

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: December 24, 2002, 16:29:47 ; Search time 1379.08 Seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724A-1

Perfect score: 1245

Sequence: 1 atgggaccctgtcagtggt.....ccaggaaagcggtgtagtag 1245

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	187.4	15.1	557	17	AQ195794
C 2	170.6	13.7	662	17	AG156539
C 3	138.2	11.1	4762	11	BC016151
C 4	137.4	11.0	477	14	BM692143
C 5	131	10.5	331	17	AQ477339
C 6	131	10.5	474	17	AQ476865

C	7	129.4	10.4	455	17	AQ476705
	8	106.8	8.6	367	9	AL709461
	9	106.8	8.6	816	12	EG540272
	10	106.2	8.5	377	14	T84802
C	11	97.8	7.9	275	12	BF990677
C	12	92.6	7.4	257	12	BF990672
	13	87.2	7.0	570	17	B89157
	14	86.4	6.9	559	12	BG722262
	15	85.6	6.9	587	17	AQ008399
	16	85	6.8	360	9	AI274154
C	17	82.8	6.7	171	14	N45658
	18	82.2	6.6	470	9	AI982914
	19	77.8	6.2	488	17	B51476
	20	75.6	6.1	285	14	T05398
	21	74.4	6.0	372	14	U46290
	22	74	5.9	845	13	BI561035
	23	70.4	5.7	301	12	BF170089
	24	70.4	5.7	678	13	BI462516
	25	67.6	5.4	487	9	AA991438
	26	62.4	5.0	392	17	AQ012375
	27	62.2	5.0	445	17	AQ145842
	28	59	4.7	544	12	BG707425
	29	57.2	4.6	413	9	AA021398
C	30	54.8	4.4	330	10	AW138646
C	31	52.2	4.2	844	17	CNS0052P
C	32	51	4.1	361	17	AQ379989
C	33	51	4.1	378	17	AQ122964
	34	50.2	4.0	734	17	BH839075
	35	50	4.0	635	17	CNS04A1A
C	36	45.6	3.7	925	17	CNS0091P
	37	44.8	3.6	839	17	CNS004NB
	38	44.8	3.6	1039	17	AG044494
	39	43.8	3.5	844	17	CNS0052P
	40	43.6	3.5	572	13	BI548831
C	41	43.4	3.5	489	17	BH638395
	42	43.2	3.5	572	12	BF616922
	43	43	3.5	925	17	CNS0091P
	44	43	3.5	1009	17	CNS010EW
	45	42.6	3.4	935	17	CNS006XK

ALIGNMENTS

RESULT 1
AQ195794/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AQ195794
RPC111-60118.TJ
RPCI-11
Homo sapiens
genomic clone
RPCI-11-60118,
DNA sequence.
GI:3607406
GI:3607406
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 557)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Other GSSs: RPC111-60118.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

passed the following selection criteria: Hexamer frequency ORF analysis
This clone has the following problem: incomplete processing.

FEATURES
source

```
1. 4762
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:313932"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT 1461 a 1022 c 1001 g 1278 t
ORIGIN
```

```
Query Match 11.1%; Score 138.2; DB 11; Length 4762;
Best Local Similarity 80.9%; Pred. No. 1.2e-24;
Matches 161; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 714 TCAATTTTGAAGAATCTTTGTCCACCCACACACCCAAAGAAATAATAACAGGA 773
Db 634 TCTAATCTGAAGAATATGTGTCCACCCACAGCTCCAATTTGAATAATAACAGGA 693

QY 774 GAGGAGGATGAATTTGGCTCTACCCCTCCAGTAGCAGAACACCTGTACCATC 833
Db 694 GAGGAGGATGAATTTGGCTCTACCCCTCCAGTAGCAGAACACCTGTACCATC 753

QY 834 TCCTTCAGTAACAGAAATAGAGACCCCTGCAAGAAATTCGCGGACTGCTACCATAGC 893
Db 754 TCCTTTGGTAGAGAAATAGAACCCCAATACAAAGAAATTTATGCTCTGCTGCCATAGC 813

QY 894 TGGAGAGCCCTTAGGACAT 912
Db 814 TGGAGAGCCCTTAGGACCT 832
```

```
RESULT 4
BM692143 477 bp mRNA linear EST 28-FEB-2002
LOCUS UI-E-CL1-aeo-f-01-0-UI-r1 UI-E-CL1 Homo sapiens cDNA clone
DEFINITION UI-E-CL1-aeo-f-01-0-UI 5', mRNA sequence.
ACCESSION BM692143
VERSION BM692143.1 GI:19005401
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS Normalization and subtraction: two approaches to facilitate gene
TITLE discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
```

FEATURES
source

```
1. 477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-CL1-aeo-f-01-0-UI"
```

```
/clone_lib="UI-E-CL1"
/tissue_type="human retina"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-CL1 is a normalized cDNA library containing the
following tissue(s): retina. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CCGCG. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

```
BASE COUNT 147 a 112 c 90 g 128 t
ORIGIN

Query Match 11.0%; Score 137.4; DB 14; Length 477;
Best Local Similarity 79.4%; Pred. No. 8e-25;
Matches 177; Conservative 0; Mismatches 41; Indels 5; Gaps 1;

QY 707 CTGAATCTCAAAATTTTGAAGAATCTTTGTCCACCCACACACCCAAAGAAATAATA 766
Db 258 CTGAATCTCAAAATCTGAAGAATATGTGTCCACCCACAGCTCCAATTTGAATAATA 317

QY 767 ACAGAGAGGAGGATGAATAATTTGGCTCTACCCCTCCAGTAGCAGAACACCTG 826
Db 318 AACAGA-----GAGGATAAAATTTGGCTCTATACCGCTCTCTCCAAATTCGAGAACATCTG 372

QY 827 TACCATCTCTTCAGTAACAGAAATAGAGACCCCTGCAAGAAATTCGCGGACTGCTA 886
Db 373 TACTGCTCTCTTCAGTAGCAGAAATAGAACCCCAATACAAAGAAATTTTACGCTCTGCTG 432

QY 887 CCATAGCTGGAGAGCCCTTAGGACATTTGCATTTTCATCTATTTTC 929
Db 433 TCATAGCTGGAGAGCCCTTAGGACCTTGGCTCTCTCTCTCTTTC 475
```

```
RESULT 5
AQ477339/c 331 bp DNA linear GSS 23-APR-1999
LOCUS CITBI-E1-2582G.TR CITBI-E1 Homo sapiens genomic clone 2582G9, DNA
DEFINITION sequence.
ACCESSION AQ477339
VERSION AQ477339.1 GI:4659458
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 331)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
JOURNAL Map Building
COMMENT Unpublished (1997)
Contact: CITBI-E1-2582G9.TF
Other_GSSs: CITBI-E1-2582G9.TF
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
```

4

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:29:47 ; Search time 1107.7 seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724A-3_COPY_1_1000

Perfect score: 1000

Sequence: 1 agccagactaggagtgcgc.....cacacataggtccagagga 1000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135.8	13.6	473	9	AL045618
2	86.2	8.6	384	17	AQ277007
3	81.4	8.1	271	17	AQ550173
4	56.6	5.7	883	12	BG719085
5	53.6	5.4	456	17	AQ307435
6	53.6	5.4	874	13	BM453647

ALIGNMENTS

RESULT 1
AL045618
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL045618 473 bp mRNA linear EST 29-FEB-2000
DKFZp434O035_r1 434 (synonym: htes3) Homo sapiens cDNA clone
AL045618
AL045618.1 GI:5433741
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 473)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
sl sequence also available.
This clone (DKFZp434O035) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

B17376 345624.TPB
BG722549 602694516
BQ261876 f274c09.y
AL181088 Tetraodon
BQ263071 f293g02.y
BQ263071 f293g02.x
C17874 C17874 Huma
BM053816 1g42h07.y
BM053916 1g44b01.y
BB636977 B8636977
BI965182 1g34f08.y
BF321581 uz63c03.y
BB666183 BB666183
BM944288 UT-M-EH0p
BQ960809 AGENCOURT
BQ284460 f290d03.y
AL270002 Tetraodon
BF800030 MRI-CI002
AL176150 Tetraodon
AL056652 Drosophil
AL298963 Tetraodon
AV392106 AV392106
AZ684627 ENTJX36TF
AL053013 Drosophil
BB730733 BB730733
BM543494 AGENCOURT
BE392622 601307451
BH313396 CH230-101
BH271688 CH230-34N
AL254287 Tetraodon
A1811650 tw44c07.x
BF192016 243073 MA
BF585113 602101835
BI151722 602916278
AL234371 Tetraodon
BQ734845 AGENCOURT
BM543930 AGENCOURT
AA562971 vl57d11.r
AQ373830 RC111-14

7 52.8 5.3 516 17 B17376
8 41.6 4.2 735 12 BG722549
9 41 4.1 364 14 BQ261876
10 39.8 4.0 1086 17 CNS02693
11 39.2 3.9 658 14 BQ263071
12 39.2 3.9 706 14 BQ262917
13 38.6 3.9 388 14 C17874
14 37.8 3.8 544 13 BM053816
15 37.8 3.8 550 13 BM053916
16 37.8 3.8 603 10 BB636977
17 37.8 3.8 627 13 BI965182
18 37.8 3.8 634 12 BF321581
19 37.8 3.8 664 10 BB666183
20 37.8 3.8 737 14 BM944288
21 37.8 3.8 942 14 BQ960809
22 37.6 3.8 619 14 BQ284460
23 37.2 3.7 993 17 CNS041BT
24 37 3.7 263 12 BF800030
25 37 3.7 1011 17 CNS020WT
26 36.8 3.7 844 17 CNS0052P
27 36.2 3.6 918 17 CNS04NOA
28 36 3.6 383 10 AV392106
29 36 3.6 905 17 AZ684627
30 36 3.6 925 17 CNS0091P
31 35.8 3.6 450 10 BB730733
32 35.8 3.6 1016 13 BM543494
33 35.6 3.6 474 10 BE392622
34 35.6 3.6 536 17 BH313396
35 35.6 3.6 728 17 BH271688
36 35.6 3.6 988 17 CNS03P7A
37 35.4 3.5 451 9 A1811650
38 35.4 3.5 568 12 BF192016
39 35.4 3.5 587 12 BF585113
40 35.4 3.5 668 13 BI151722
41 35.4 3.5 818 17 CNS03902
42 35.4 3.5 846 14 BQ734845
43 35.4 3.5 1280 13 BM543930
44 35.2 3.5 497 9 AA562971
45 35.2 3.5 532 17 AQ373830

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE AUTHORS

1 (bases 1 to 1086)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.

TITLE

Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL AUTHORS

Unpublished
2 (bases 1 to 1086)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL AUTHORS

Unpublished
3 (bases 1 to 1086)
Genoscope.
Submitted (12-APR-2000)
Direct Submission

TITLE

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES

Location/Qualifiers
1..1086
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="240B09"
/clone_lib="G"
/note="Genoscope sequence ID : COAG240CA05SP1-end : PUC-ori"

BASE COUNT

286 a 222 c 284 g 260 t 34 others

ORIGIN

Query Match 4.0%; Score 39.8; DB 17; Length 1086;
Best Local Similarity 53.3%; Pred. No. 96;
Matches 57; Conservative 11; Mismatches 39; Indels 0; Gaps 0;

QY 888 TGGACACAGCTGGCACTGCTCAACAGCGCTACCAAGACATCATGTTTTTTTTTT 947
DB 1022 TTGAACAGAGCTMBCCSCCTTAAAGCGYGCAGAAATCTTTTTTTTTTTKTYTC 963

QY 948 TCCACAAACCTGGACCTGAATGGGTGTGGACACACATAGAGTCC 994
DB 962 TCCCCCCCACAGGCCCTCCAGGGGGAGCGGTCCGCCAGRAYTCC 916

RESULT 11

BO263071/c
LOCUS BO263071 658 bp mRNA linear EST 06-MAY-2002
DEFINITION fz39g02.y1 Sugano SUD adult male Danio rerio cDNA clone 5916098 5' similar to TR:Q9WTU0 Q9WTU0 PHD-FINGER PROTEIN. ; mRNA sequence.
ACCESSION BO263071
VERSION BO263071.1 GI:20463850
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
1 (bases 1 to 658)

REFERENCE AUTHORS

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Willson, R.

TITLE

WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA Sequencing by Washington University Genome Sequencing Center Clone distribution information can be found through the I.M.A.G.E. Consortium/LINL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 526.

FEATURES

Location/Qualifiers
1..658
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5916098"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="vector: pME18S-FL3; Site 1: DraIII (CACCATGTG); Site 2: DraIII (CAGTGTGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGGGCCCTTTTITTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CAGTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCAC."

BASE COUNT

169 a 131 c 173 g 185 t

ORIGIN

Query Match 3.9%; Score 39.2; DB 14; Length 658;
Best Local Similarity 62.0%; Pred. No. 1.4e+02;
Matches 62; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 316 CACTAATAGTTTGGTGGGCTCTTTCTGGAAGCTACCTCTCTCTGTTGCCCTC 375
DB 333 CACTATAGGGTAATGAAGCCCTCTTTCTCAAGAAGCGTTTGGTCACTGATCCCTG 334

QY 376 CATCTTCCCCAACCCAGTACTTCTGGCCATCCTCTCTGTC 415
DB 333 CATCTGCACAGATATACCGCTGGCAAGCTCTCTGCC 294

RESULT 12

BO262917
LOCUS BO262917 706 bp mRNA linear EST 06-MAY-2002
DEFINITION fz39g02.x1 Sugano SJD adult male Danio rerio cDNA clone 5916098 3' similar to TR:Q9WTU0 Q9WTU0 PHD-FINGER PROTEIN. ; mRNA sequence.
ACCESSION BO262917
VERSION BO262917.1 GI:20463696
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
1 (bases 1 to 706)

REFERENCE AUTHORS

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Willson, R.

TITLE

WashU Zebrafish EST Project 1998
Unpublished (1998)
Other_ESTs: fz39g02.y1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:29:47 ; Search time 1108.8 Seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724A-3_COPY_50000_51000
Perfect score: 1001
Sequence: 1 agcaactgtaagtccggc.....ggccctgctgcatgtgacc 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.6	15.4	392	12	BF904043
2	135	13.5	331	12	BF901889
3	132	13.2	460	17	AQ478007
4	132	13.2	480	17	AQ198951
5	132	13.2	542	17	AQ630303
6	132	13.2	613	17	AQ427668

7	132	13.2	701	17	AQ473751
8	131	13.1	502	17	AQ692634
9	125.8	12.6	524	17	AQ440312
10	123	12.3	509	17	AQ757386
11	116.2	11.6	430	17	AQ494621
12	110.2	11.0	650	17	AG176043
13	105.2	10.5	393	12	BF901864
14	105.2	10.5	497	17	AQ169608
15	104.8	10.5	578	17	AQ008424
16	103.4	10.3	409	17	AQ674424
17	91.8	9.2	325	17	AQ197958
18	85.8	8.6	456	17	AQ696533
19	85.2	8.5	525	17	AQ784251
20	83.6	8.4	490	17	AQ626785
21	80.4	8.0	404	17	AQ086415
22	80	8.0	330	17	AQ817144
23	79.4	7.9	623	17	AZ521679
24	78.2	7.8	535	17	AQ879443
25	77.2	7.7	501	17	AQ611000
26	69.6	7.0	222	17	B88266
27	61.2	6.1	711	10	BE566324
28	52.4	5.2	669	17	AG034460
29	50.6	5.1	685	17	AQ312269
30	50.4	5.0	718	17	B92472
31	50.2	5.0	689	17	AG184567
32	49	4.9	1007	14	BM806213
33	48.8	4.9	707	17	AQ351078
34	48.4	4.8	1063	14	BM920247
35	47.2	4.7	395	12	BF901989
36	47	4.7	803	14	BM983069
37	46.4	4.6	540	17	AQ756822
38	46.4	4.6	852	17	AQ746055
39	46.2	4.6	663	17	AG089297
40	46	4.6	502	17	AQ155233
41	45.8	4.6	401	17	AQ032208
42	45.6	4.6	366	12	BF901856
43	45.6	4.6	567	17	AQ601940
44	45.4	4.5	440	17	AQ130549
45	45	4.5	752	17	AQ939201

ALIGNMENTS

RESULT 1
BF904043
LOCUS BF904043 392 bp mRNA linear EST 18-JAN-2001
DEFINITION IL5-MT0210-201200-338-g09 MT0210 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF904043
VERSION BF904043.1 GI:12295502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-MT0210-201200-338-g09&t3=2000-12-20&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 382.
Location/Qualifiers

FEATURES

source
1. .392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0210"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
88 a 118 c 111 g 75 t

Query Match 15.4%; Score 154.6; DB 12; Length 392;
Best Local Similarity 68.4%; Pred. No. 2.2e-35;
Matches 281; Conservative 0; Mismatches 109; Indels 21; Gaps 4;
QY 392 GGAGGAAGGTGAGCAGCTGGCGCTGCTGCTTTGAGGGAGGATGGCGATGTGAAAGTCAGT 451
Db 1 GGAGCAAGGTGAGCAGCTGGCGCACTACGTTGGATGAGGGTGGGATGTGGACGTACC 60
QY 452 GACCACCGTGGGAGGACACTCCCTGGCTCCATCTCTGATCTTTAGATTATTGGGACA 511
Db 61 GCCACCCCTGGCAGGACACTCCCTGGCTCCATCTCTACATCTTAGCTTACTGGGAGT 120
QY 512 GTTTCATACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 571
Db 121 GTTTCACACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 178
QY 572 ATCAATGATAAATTCCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631
Db 179 ATCAAGACCRAATTCCTGAGGAGGGGCTGTTTATATCAACTCTGAGAACAGGCTG--- 235
QY 632 CTACATGGGATTTGAGGGAGGGTGGAGCCCTTAAAGAAAGCCCGACAGACTGCCCC 691
Db 236 -----GGGGCGCTGCATGGGACCCCTTCAGAAAACGTGCCCGAGGACGAGCCC 283
QY 692 TGCCTCTCTCTCCCCCAAGTTCATTTATTTATCTTCCACCCAGGAGCTGTCAGATC 751
Db 284 ---CTTCTCCTGCTGCCAAAGGCCCCCACAGCAACTTCCACCCAGGCGCTGTCAGCATC 341
QY 752 CTG---CCCTTCCTCTCCAGATCAAGTCTTTCAGGAATCAACTACTTTC 800
Db 342 CTGCTCTCTCTTAGGAGGTCGAGTTCGAGGAATGAGCTGCTCC 392

RESULT 2
BF901889 331 bp mRNA linear EST 18-JAN-2001
LOCUS IL5-MT0210-111200-337-g04 MT0210 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF901889
ACCESSION BF901889.1 GI:12293348
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 331)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
MEDLINE
COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-MT0210-111200-337-g04&t3=2000-12-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 330.
Location/Qualifiers

FEATURES

source
1. .331
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0210"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
81 a 102 c 93 g 55 t

Query Match 13.5%; Score 135; DB 12; Length 331;
Best Local Similarity 75.9%; Pred. No. 1.7e-29;
Matches 180; Conservative 0; Mismatches 55; Indels 1; Gaps 1;
QY 392 GGAGGAAGGTGAGCAGCTGGCGCTGCTGCTTTGAGGGAGGATGGCGATGTGAAAGTCAGT 451
Db 1 GGAGCAAGGTGAGCAGCTGGCGCACTACGTTGGATGAGGGTGGGATGTGGACGTACA 60
QY 452 GACCACCGTGGGAGGACACTCCCTGGCTCCATCTCTGATCTTTAGATTATTGGGACA 511
Db 61 GCCACCCCTGGCAGGACACTCCCTGGCTCCATCTCTACATCTCAGGCTTACTGGGAGT 120
QY 512 GTTTCATACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 571
Db 121 GTTTCACACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 178
QY 572 ATCAATGATAAATTCCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 628
Db 179 ATCAAGACCRAATTCCTGAGGAGGGGCTGTTTATATCAACTCTGAGAACAGGCTG 235

RESULT 3
AQ478007 460 bp DNA linear GSS 23-APR-1999
LOCUS CITBI-E1-2585E13.TR CITBI-E1 Homo sapiens genomic clone 2585E13,
DEFINITION AQ478007
ACCESSION AQ478007
VERSION AQ478007.1 GI:4660126
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building


```
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
source
1. .542
/organism="Homo sapiens"
/db_xref="GDB:7681049"
/db_xref="taxon:9606"
/clone="RPCI-11-472H18"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 143 a 137 c 116 g 146 t
ORIGIN

Query Match 13.2%; Score 132; DB 17; Length 542;
Best Local Similarity 85.9%; Pred. No. 1.8e-28;
Matches 158; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

Qy 1 AGCAACCTGTAAGTTGGGCTGCAATCATAGATAAGTAAGTGGAGCTTGTATGGGCAG 60
Db 17 AGCAGCTTGGAAATTTGGCTGCAATCATACAAAGAGAGCTTGTATGGGCAG 76

Qy 61 GGATGGCTGCAGCTTCATGGATAGAAATGTCACAGTTGGGCTAGATACCAACAT-GG 119
Db 77 GGATGCTGCAGCTGCAGCTGGGTAGAAATGACCACTTGGGCCAGACATATCCAAACATGG 136

Qy 120 GGGCTCCACTCCTCTTTGTAGCACAGCACCACATAGAAAGAGATAAGCAACTTGGAGTAG 179
Db 137 GGGCCCCACCCCTCTTTGTAGCATATGCACAGTAGAAAAGAGATAAGCAACTTGGAGTAG 196

Qy 180 CTCA 183
Db 197 CTCA 200

RESULT 6
LOCUS AQ427668 613 bp DNA linear GSS 24-MAR-1999
DEFINITION CITBI-El-2575C21.TR CITBI-El Homo sapiens genomic clone 2575C21,
DNA sequence.
ACCESSION AQ427668
VERSION AQ427668.1 GI:4500575
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 613)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
source
1. .613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2575C21"

FEATURES
source
1. .542
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2575C21"
/clone_lib="CITBI-El"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 200 a 173 c 146 g 181 t
ORIGIN

Query Match 13.2%; Score 132; DB 17; Length 701;
Best Local Similarity 85.9%; Pred. No. 2e-28;
Matches 158; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

Qy 1 AGCAACCTGTAAGTTGGGCTGCAATCATAGATAAGTAAGTGGAGCTTGTATGGGCAG 60
Db 209 AGCAGCTTGGAAATTTGGCTGCAATCATACAAAGAGAGCTTGTATGGGCAG 268

Qy 61 GGATGGCTGCAGCTTCATGGATAGAAATGTCACAGTTGGGCTAGATACCAACAT-GG 119
Db 269 GGATGCTGCAGCTGCAGCTGGGTAGAAATGACCACTTGGGCCAGACATATCCAAACATGG 328

Qy 120 GGGCTCCACTCCTCTTTGTAGCACAGCACCACATAGAAAGAGATAAGCAACTTGGAGTAG 179
Db 329 GGGCCCCACCCCTCTTTGTAGCATATGCACAGTAGAAAAGAGATAAGCAACTTGGAGTAG 388

Qy 180 CTCA 183
Db 389 CTCA 392

RESULT 7
LOCUS AQ473751 701 bp DNA linear GSS 23-APR-1999
DEFINITION CITBI-El-2585N20.TF CITBI-El Homo sapiens genomic clone 2585N20,
DNA sequence.
ACCESSION AQ473751
VERSION AQ473751.1 GI:4652012
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
Location/Qualifiers
source
1. .701
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2585N20"
/clone_lib="CITBI-El"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 200 a 173 c 146 g 181 t
ORIGIN

Query Match 13.2%; Score 132; DB 17; Length 701;
```


Best Local Similarity 64.7%; Pred. No. 1.4e-26; Matches 187; Conservative 0; Mismatches 102; Indels 0; Gaps 0;																	
Qy	392	GGAGGAAGTGGACAGCTGGCGCTGTGCGTTTGGAGGAGGATGGCGATGTGAAAGTCAGT	451														
Db	159	GGTGTAGTCCACATGACAGAGGGGTGGGTGGGTATTCATCGGCCAGGGTTTACT	218														
Qy	452	GACCACCGTGGGAGGACACTCCCTGGCTCCATCCTCTGCATCTTAGATTATTGGGACA	511														
Db	219	GGATTCCCTGGGACGACACTGCTGCTCTATTCTCTTCATCTTAGATGATTGGGAGG	278														
Qy	512	GTTTGATACAGAGAAGGAGAGACCCATCCCAATGGAGGGTTTGTATAGATGAATATA	571														
Db	279	GTATGATACAGAGAAGGAGAGACCCATTTCATATAAGGGTGTGAATAGGAAACTGA	338														
Qy	572	ATCAATGATAAATCTCTAGAGGAGGACATTTTATAATCAACTCTGAGAACAGGTTGGAG	631														
Db	339	AATCAAGGACCAATCTTTCGAGAAGGCTGCTGTGATCTAATCTGAGAACACACTGGGG	398														
Qy	632	CTACATGGGATTGGAGGGAGGGTGGAGCCCTTAAAAAGAAAGCCCA	680														
Db	399	CGGCATGGGATTGGAGGAGGAGGAGTCTCTTTAGAGAAAGCTCTCA	447														
RESULT 10																	
LOCUS		A0757386		509 bp		DNA		linear									
DEFINITION		HS_5463_B2_E10_T7A RPCI-11 Human Male BAC Library Homo sapiens		genomic clone Plate=1039 Col=20 Row=J, DNA sequence.													
ACCESSION		A0757386															
VERSION		A0757386.1		GI:5622256													
KEYWORDS		GSS.															
SOURCE		human.															
ORGANISM		Homo sapiens															
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.															
AUTHORS		Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.															
TITLE		Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome															
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)															
MEDLINE		99380589															
COMMENT		Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 1039 row: J column: 20 Seq primer: T7 Class: BAC ends High quality sequence stop: 509.															
FEATURES		source															
		1..509		Location/Qualifiers													
		/organism="Homo sapiens"															
		/db_xref="taxon:9606"															
		/clone="Plate=1039 Col=20 Row=J"															
		/clone_lib="RPCI-11 Human Male BAC Library"															
		/sex="male"															
		/note="Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"															
BASE COUNT 134 a 125 c 109 g 138 t 3 others																	
ORIGIN																	
Query Match 12.3%; Score 123; DB 17; Length 509; Best Local Similarity 82.6%; Pred. No. 9.3e-26; Matches 152; Conservative 0; Mismatches 31; Indels 1; Gaps 1;																	
Qy	1	AGCAACCTGTAAGTTCGGGCTGCAATCATAGATAAGTAAGTAGGAAGCTTTATGGGCGAG	60														
Db	48	AGCAGCCTGGANAATTGGGCTGCAATCATACAAAAGGAAGCTGGAAGTCTTTTGGGTAG	107														
Qy	61	GGATGGCTGCAGCTTCATGGATAGAAATGTCACGCTGGGCTAGATACATCAACAT	119														
Db	108	GGATGCCTGCAGCTGCATGGGTAGAAATGCACCACTTGGGCCACACATATCCACATGGG	167														
Qy	120	GGGCTCCTACTCTCTTTGTAGCACAGCACCACATAGGAAGAGATAAGCAACTTGGAGTAG	179														
Db	168	GGGCCCAACCCCTCTTTGTAGCATATGCACAGTAGAAAAGAGATAAGCAACTTGGAGTAG	227														
Qy	180	CTCA 183															
Db	228	CTCA 231															
RESULT 11																	
LOCUS		A0494621		430 bp		DNA		linear									
DEFINITION		HS_5171_B1_H10.SP6E RPCI-11 Human Male BAC Library Homo sapiens		genomic clone Plate=747 Col=19 Row=P, DNA sequence.													
ACCESSION		A0494621															
VERSION		A0494621.1		GI:4694788													
KEYWORDS		GSS.															
SOURCE		human.															
ORGANISM		Homo sapiens															
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.															
AUTHORS		Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.															
TITLE		Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome															
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)															
MEDLINE		99380589															
COMMENT		Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 747 row: P column: 19 Seq primer: SP6 Class: BAC ends High quality sequence stop: 430.															
FEATURES		source															
		1..430		Location/Qualifiers													
		/organism="Homo sapiens"															
		/db_xref="taxon:9606"															
		/clone="Plate=747 Col=19 Row=P"															
		/clone_lib="RPCI-11 Human Male BAC Library"															
		/sex="male"															
		/note="Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"															
BASE COUNT 134 a 125 c 109 g 138 t 3 others																	
ORIGIN																	
Query Match 12.3%; Score 123; DB 17; Length 509; Best Local Similarity 82.6%; Pred. No. 9.3e-26; Matches 152; Conservative 0; Mismatches 31; Indels 1; Gaps 1;																	

```
BASE COUNT      129 a      97 c      98 g      105 t      1 others
ORIGIN

Query Match      11.6%; Score 116.2; DB 17; Length 430;
Best Local Similarity 78.1%; Pred. No. 9.6e-24;
Matches 139; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 6 CCTTAAGTTTCGGGCTGCAATCATAGATAAGTAAGATGGAAGCTTGATGGCAGGGATG 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 CCCGCAAAATTCGCTGTAATCATACAAAGGAAGCTGGAAGCTTGCTGGCGAGGGATG 93

Qy 66 GCTGAGCTTCATGATAGAAATGTCAGCTTGGGCTGATACATACCAACTGGGGGCTC 125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94 CTTGCACCTGTCATGGTAAAAATGACCACTTGGGCCAGACATATACAACTGGGGGCC 153

Qy 126 CACTCCTCTTTGTAGCACACCACCATAGGAAGAGATGAACCAACTGGAGTAGCTCA 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 CNACCCCTCTTTGGTACATATGCAGTAGAAGAAAGAGATAAGCAACTTGGAGTAGCTCA 211

RESULT 12
AG176043
LOCUS      AG176043          650 bp      DNA      linear      GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-047C16.TJ, genomic survey
ACCESSION      AG176043
VERSION      AG176043.1 GI:16705723
KEYWORDS      GSS.
SOURCE      Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-047C16.TJ.
ORGANISM      Pan troglodytes
REFERENCE      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library RPCI-43
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 650)
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbos@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT      Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
PRIMERS

Sequencing: TJ
LIBRARY
Vector      : pBACe3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI
Location/Qualifiers
1..650
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-047C16.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      150 a      162 c      155 g      183 t
ORIGIN

Query Match      11.0%; Score 110.2; DB 17; Length 650;
Best Local Similarity 79.8%; Pred. No. 8.1e-22;
Matches 130; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 839 AATTGGGGTTTGGTCCCGATGCCAGTCAAGTGCACAGTCAGATAAAGGTCAGAGCTTA 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 11 ATTCGGGGTTTGGTCCCGTCTCTGAAGTGTGCACAGTCAGATAAAGGTCAGGGCCTA 70

Qy 899 GGAGATTAGCGAGGAGTAGAAGAACTCTCTCTGTGTACACAGCTTCAGAGAGCTGGG 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 ACAGATTAGCAGATGGTAGGAAGACTCTATCTTGCAGCCAGCTTCAGAGAGCTGTGG 130

Qy 959 CCATGGCTTCTCGTGTCAACATTAGGCCCTCGCTGCATGGTGACC 1001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 CCATGGCTCCCGAGTCAACATTAGGCCCTGTTCCTGGGAACC 173

RESULT 13
BF901864
LOCUS      BF901864          393 bp      mRNA      linear      EST 18-JAN-2001
DEFINITION IL5-WT0210-111200-337-h09 MT0210 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF901864
VERSION      BF901864.1 GI:12293323
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 393)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5et2-IL5-WT0210-
            111200-337-h09&t3=2000-12-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 392.
FEATURES
            source
            1..393
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="MT0210"
            /dev_stage="Adult"
            /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
            SmaI; A mini-library was made by cloning products derived
            from ORESTES PCR (U.S. Letters Patent application No. 196
            716 - Ludwig Institute for Cancer Research) profiles
            into the pUC 18 vector. Reverse transcription of tissue
            mRNA and cDNA amplification were performed under low
            stringency conditions."
BASE COUNT      97 a      115 c      98 g      83 t
ORIGIN

Query Match      10.5%; Score 105.2; DB 12; Length 393;
Best Local Similarity 62.4%; Pred. No. 1.1e-20;
Matches 257; Conservative 0; Mismatches 133; Indels 22; Gaps 5;

Qy 468 ACACCTCCCTGGCTCCATCCTCTGCATCTTAGATTATTTGGGACAGTTTGTACACAGAGA 527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACATCCGTGGATCCATCCTCTACATCGTAGGCTTACTGGAGTGTTTGTACACACACAGG 60

Qy 528 AGGAGGAGACCCATCCCAATGGAGGGTTTGATTAGATGAATATATCAATCAATAAATTC 587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 61 AGGAGGGACCTATCCCA--GGAGAAAATGGGAAAGCACTGGAAATCAAGACCAATTC 118
Qy 588 TAGAGAGGGAGCTTTTATATCAACTCTGAGACAGCTTGGAGCTACATGGGATTGGAG 647
Db 119 TCGAGAGGGCTGTTATATCAACTCTGAGACAGCTG-----GGGG 163
Qy 648 GGGAGGGTGGAGCCCTTAAAGAAAAGCCCGAGAGACTGCCCTGCGCTCTCTCTCC 707
Db 164 CGCTGATGGGACCCCTTCAGAAAATGCCCCAGGACAGCCC--CTTCTCCCTGCGCTG 221
Qy 708 CACAGTTCATTTATATCTCCACCCAGGAGCTGTCCAGATCCG--CCCTCCGCTCT 765
Db 222 CAAAGGCCCCCAAGCACTTCCACCCAGGCCCTGTCCAGCATCTGCTCTGTCTCTA 281
Qy 766 CCAGATCAAGTCCCTTCAGGAAATGCACTACTTC-AGTGACAAGAGATAATTCATCT 824
Db 282 GGAGTCCGAGTCTGCGAGGAATGCTCTCAAGTGGCTGCCATGAATTACAGCT 341
Qy 825 TCTGACAGAGGAGGAATTTGGGGTTGTCCCGAGTCCCATGAAGTGGCACAGT 876
Db 342 TCGTCTCTTGAGAAATTTGTACCTATTTCCAAAGAAATGGAGCAGCTCAGT 393

RESULT 14
Aq169608 497 bp DNA linear GSS 16-OCT-1998
LOCUS HS_3176_B2_G12_T7 CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=3176 Col=24 Row=N, DNA sequence.
ACCESSION Aq169608
VERSION Aq169608.1 GI:3567283
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 497)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3176 row: N column: 24
Class: BAC ends
High quality sequence stop: 497.
FEATURES
source
Location/Qualifiers
1..497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3176 Col=24 Row=N"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 114 a 153 c 82 g 148 t
ORIGIN
Query Match 10.5%; Score 105.2; DB 17; Length 497;
Best Local Similarity 84.9%; Pred. No. 2.2e-20;
Matches 141; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

Qy 17 GGGCTGCAATCATAGTAAGTGAAGCTTCTATGGCC-AGGGATGGCTGCAGCTT 75
Db 48 GTGCTGCAATCATAGAAAAGAGCTTGTGTGGGCTAGGAGTGGCTGCAGCTG 107
```

```
Qy 76 CATGATAGAAAATGTCAGCTTGGGCTAGATACATCAACAT--GGGGGCTCCACTCTCT 134
Db 108 CATGGTAGAAAATGACCACTTGGCCAGACATATCAACATGGGGGGCCCACTCTCT 167
Qy 135 TTGTAGCACAGCCACCATAGGAAGAGATAAGCAACTTGGAGTAGC 180
Db 168 TTGTAGCATACGCACAGTATATAAGAGATAAGCACTTGGAGTAGTATC 213

RESULT 15
Aq008424 578 bp DNA linear GSS 27-JUN-1998
LOCUS CIT-HFP-2062M17.TF CIT-HSP Homo sapiens genomic clone 2062M17, DNA
DEFINITION sequence.
ACCESSION Aq008424
VERSION Aq008424.1 GI:3128531
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 578)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden
,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..578
/organism="Homo sapiens"
/db_xref="GDB:7061300"
/db_xref="taxon:9606"
/clone="2062M17"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 148 a 146 c 129 g 155 t
ORIGIN
Query Match 10.5%; Score 104.8; DB 17; Length 578;
Best Local Similarity 87.5%; Pred. No. 3.2e-20;
Matches 126; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

Qy 41 ATGGAAGCTTGTATGGGAGGATGGCTGCAGCTTCATGATAGATAAATGTCAGCTTGGG 100
Db 43 ATGCAAGCTTGTGGGAGGATGCCCTGCAGCTGCATGGGTGAGAAATGACCACTTGGG 102
Qy 101 CTAGATACATCAACAT--GGGGGCTCCACTCTCTTTTGTAGCACAGCAGCATAGGAAAG 159
Db 103 CCAGACATATCCACATGGGGGGCCCCACCCCTCTTTGTAGCATATGCACATAGAAAG 162
Qy 160 AGATAAGCAACTTGGAGTAGCTCA 183
Db 163 AGATAAGCAACTTGGAGTAGCTCA 186

Search completed: December 24, 2002, 21:54:01
```


Job time : 1115.8 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:29:47 ; Search time 1108.8 Seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724a-3_COPY_70000_71000

Perfect score: 1001

Sequence: 1 ggagatggataaacgctgtg.....ccaatcaggagtctatgtg 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	522.2	52.2	640	12	BG534918 602554106
C 2	418.6	41.8	808	9	AU136659 AU136659
C 3	357.8	35.7	649	10	BE150630 RC3-HT027
C 4	307.6	30.7	524	12	BG003846 MR3-GN018
C 5	299	29.9	507	12	BF879658 RC2-ET018
C 6	281.8	28.2	500	12	BF879654 RC2-ET018

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	280.6	28.0	625	13	BI046524
8	272	27.2	464	13	BI057039
9	250.6	25.0	658	10	AV721454
10	248.2	24.8	743	17	AG183922
11	237.2	23.7	574	12	BG77821
12	223.8	22.4	636	17	AQ00691
13	217.8	21.8	418	9	AI830956
14	217.8	21.8	524	12	BF725513
15	217.8	21.8	576	13	BI115904
16	217.8	21.8	579	10	BE263504
17	217.8	21.8	591	10	BE258753
18	217.8	21.8	598	12	BG716902
19	217.8	21.8	693	13	BI547646
20	217.8	21.8	702	12	BE795896
21	217.8	21.8	709	12	BG705551
22	217.8	21.8	738	13	BI767440
23	217.8	21.8	810	12	BG764162
24	217.8	21.8	820	14	BM725035
25	217.8	21.8	838	14	BQ217512
26	217.8	21.8	861	13	BI916025
27	217.8	21.8	881	14	BQ224422
28	217.8	21.8	888	13	BI547711
29	217.8	21.8	908	14	BQ684259
30	217.8	21.8	916	13	BI757276
31	217.8	21.8	918	12	BG760309
32	217.8	21.8	977	12	BG395462
33	217.8	21.8	978	14	BQ683756
34	217.8	21.8	1008	13	BM475273
35	217.8	21.8	1019	13	BM468970
36	217.8	21.8	1079	14	BM800060
37	217.8	21.8	1094	14	BM811369
38	217.4	21.7	511	10	AW162774
39	217	21.7	803	12	BG431001
40	217	21.7	1081	12	BE907415
41	216.8	21.7	354	13	BI057453
42	216.2	21.6	417	10	AW370537
C 43	215	21.5	446	9	AA988929
C 44	214	21.4	627	10	AV647395
C 45	210.2	21.0	430	10	AW104316

ALIGNMENTS

RESULT 1	BG534918/c	BG534918	640 bp	mRNA	linear	EST 03-APR-2001
LOCUS	602554106F1 NIH_MGC_77	Homo sapiens	cdna clone IMAGE:4663824 5',			
DEFINITION	mRNA sequence.					
ACCESSION	BG534918					
VERSION	BG534918.1	GI:13526461				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 640)					
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaphs-r@mail.nih.gov Tissue Procurement: CLONTECH Laboratories, Inc. CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI466 row: 1 column: 01 High quality sequence stop: 639.					
FEATURES	Location/Qualifiers					
source	1..640					

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:4663824"  
/clone_lib="NIH_MGC_77"  
/lab_host="DH10B (TI phage-resistant)"  
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:  
SfiI (ggcgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCCGCGCGGCGACATG-dh(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.9  
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH_MGC Library."  
BASE COUNT      170 a 134 c 166 g 170 t  
ORIGIN  
  
Query Match      52.2%; Score 522.2; DB 12; Length 640;  
Best Local Similarity 90.6%; Pred. No. 7e-152;  
Matches 579; Conservative 0; Mismatches 58; Indels 2; Gaps 2;  
  
QY 210 CAGAGCTGAGAAACTGGTGTAGTGCAGGCTCAGGCAAAACCCCTGACCTCCATGTTA 269  
Db 638 CAGGAATTGAAAAATGCTGGAGCGCAGGCTCAGGCGAAAAACCCAGACTCCATGTTA 579  
  
QY 270 TGGCCATGCTAGC-TGTAATATCTCTGCAGTATGATTTTCTGTGCAGAGCAAAACA 328  
Db 578 TGGCCATGCTAGCTGTAATATCTCTGTAGGTATGTTTCTCCTGTGCAGAGC-AATACA 520  
  
QY 329 TATTGGGCATATTTCTTAACCCACCGGTAGTGTATCATCTCTGAAGCAGCAGCTCCTC 388  
Db 519 TATTGGGCATATGTTCTTAACCCACTGCTAGTGTGACCATACTCTGGAACGACACTCCTC 460  
  
QY 389 CTGAGATATATCATGATCAAGAGCATCAGTACCAGGACCTCTAACTCCCTCCACAG 448  
Db 459 CTGAGATATATCATGATCAAGAGCATAGGCACCGACCCCTAACTCTCCCTGCACAG 400  
  
QY 449 AGCAATTAGACTCTCATAAACAATGATATCAATATACACATCCCATTTGGAGGACTTCCTT 508  
Db 399 AGCAATTAGACTCTCAGNACAATGATATCAATATACACATCCCATTTGGAGGACTTCCTT 340  
  
QY 509 TATGTGCACCCAGATATGCTCAACTGCAGTTGGCTTTGCGATTTGATFCCCAAGCAT 568  
Db 339 TATGGCTCACCCAGGATACCTCACTCACTGCTGCTTTGCAATTTCAATFCCCAAGCAT 280  
  
QY 569 GGTGTAGTTACCATAAAAAATATGTACCTATTAGACCTTTAGCTTTATTAATATTACTT 628  
Db 279 GGCTGAGTTACCATGGAATATATGTACCTATTAGGCTTTAGCTTCTATTATTACTG 220  
  
QY 629 GTGTAGTTACTAATCACTCTCCGGCCCCATCACCAAAATTTGATGATTATACGAATGGG 688  
Db 219 GTGTGTTTACTAATCACTCTCCGGCCCCATCGCCAAATTTGATTATACGAATGGG 160  
  
QY 689 CTCCTTTGATTAATCTCACCCCTCTCTGGGCCACTGCTTTGGCCCTTTAGCTAGAC 748  
Db 159 CTCCTTTGATTAATCTCACCCCTCTCTGGGCCACTGCTTTGGCCCTTTAGCTAGAC 100  
  
QY 749 AATAGTCCATGTTAATGGGAGACATATTAGTGGGGTCCCTGTGGTTCATTAAAGATGGG 808  
Db 99 AACAGTCCATGTTAATGGGAGACATATTAGTGGGGTCCCTGTGGTTCATTATGATGGG 40  
  
QY 809 GAGATGAGAATCAGACCATGGCATAAATTCACCTGGC 847  
Db 39 GAGATGAGAATCAGACCTCAGGCATAGCTCTACAGC 1
```

```
RESULT 2  
AUI36659  
LOCUS  
DEFINITION AUI36659 PLACE1 Homo sapiens cdna clone PLACE1004793 5', mRNA  
sequence.  
ACCESSION AUI36659
```

```
VERSION AUI36659.1 GI:10997198  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 808)  
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
Tsugai,T.  
TITLE HRI human cdna project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cdna project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cdna library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
FEATURES  
source Location/Qualifiers  
1..808  
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="PLACE1004793"  
/clone_lib="PLACE1"  
/tissue_type="placenta"  
/note="Vector: PME18SFL3"  
BASE COUNT 227 a 186 c 169 g 219 t 7 others  
ORIGIN  
  
Query Match 41.8%; Score 418.6; DB 9; Length 808;  
Best Local Similarity 76.1%; Pred. No. 2.2e-119;  
Matches 553; Conservative 0; Mismatches 171; Indels 3; Gaps 3;  
  
QY 257 GACTCCATGTTTATGGCCATGCTAGCTGTATATATCCCTGTCAGTATGATTTTCTGTGCA 316  
Db 1 GATTCCTGTTCTTGGCCATGTTAGCCATATATATCCCTGTCAGTATGTTTCTGTGCA 60  
  
QY 317 GAAGCAAAACATATTTGGGCATATTTTCTTAACCCACGGTAGTGTGA-TCATACTCTGA 375  
Db 61 GAGCAAAACATATTTGGGCATATTTTCCCAAGCCCCAGCAGTATGACCATACTTTGG 120  
  
QY 376 AGCAGCACTCTCTCTGAGATATATCATGATCAAGGAGCATCAGTACCAGGACCTCTAACT 435  
Db 121 AGTCACTCTCTCTAGATTTATCATGATTTAGGAGCATGGGCTCCAGGACCCCTAACT 180  
  
QY 436 CCCCTGACAGAGCAATTAGACTCTCATAAACAATGGTATCAATATPACACATCCATTG 495  
Db 181 CCACCTGACATAGAACAGTTAGACTCTCAGAATAATATGTCAATTAATATACGCTCCATTG 240  
  
QY 496 GAGGACTTCCTTTATGTGTCACCCAGGATACATGCTCAACTGCAGTTGCCCTGCAGTT 555  
Db 241 GAAGGACTTCCTTTGTGTGTCACCAAGACATCACTCAGCCCATAGCTGCTTTACAGTT 300  
  
QY 556 TGATCCCAAGCATGGTTGAGTTACCATAAAAAATTTATGTACCTATTAGACCTTAGCTTT 615  
Db 301 CAAGCTCACACATGGTTGAGTCATCTATGGGAAATCATGTAATTTAAGTCTTGGTTAT 360  
  
QY 616 ATTAATATTACTTGTGTAGTTTACTTAATCACTCTCTGGGCCCATCACCCAAATTTAGTGA 675  
Db 361 ATTAATGTAACCGGTGCTTAACCAACCATCTCTGGGCCAATCGCTTCATTGTGCTGAC 420  
  
QY 676 TATACAGATGGGCTCCCTTTTGATAATTTCTCACCCCTCTCTGGGCCCATCTGCTTGGC 735  
Db 421 TATACAGATGGATTCCTTCAATAGTTCTTACCCCTCTCCATAGACCCAGCTGCTTGGC 480  
  
QY 736 CCCTTAGTACACAAATAGTCCATGTTAATGGGAGACATATTAGTACTGGGTCCTCTGTGTT 795  
Db 481 CCATGGCTAGAAACAAATCTATGTTAACTGGAGACATTTGGATGGGGACCTAAAGGC 540
```


SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 658)
REFERENCE
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu
, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
Chen, J., Chen, Z., and Han, Z.
TITLE Homo sapiens cDNA HTB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source 1. .658
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTBAKH02"
/clone_lib="HTB"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 181 a 151 c 176 g 149 t 1 others
Query Match 25.0%; Score 250.6; DB 10; Length 658;
Best Local Similarity 78.8%; Pred. No. 7.6e-67;
Matches 339; Conservative 0; Mismatches 85; Indels 8; Gaps 3;
QY 1 GGAGATGATAAACCCTGTGAGTGCCCTCAAGTTGTGTGGACCATGGAATGGGAGACTG 60
Db 219 GGAGATGGACAACCGTGTGGTGCCCTCCAGGTGTGTGACCATGGAACGGGAGACCG 278
QY 61 GAGGATACATGATGCCAATCAAGCCGACGCTCCAGATGAGCCATGAGCCAGTT 120
Db 279 GAGGATCCATGTTATTCACACCGCTGGCCCTGTACCTCCAGTAGAGCCATGAGCCAGC 338
QY 121 GAATCTCAATGTGAAGATGGAATGAACCGCAGAGAGTACACTGACGTCACACCTCAT 180
Db 339 GAATCTCAATGTGAAGATGGAATGAACCGCAGAGAGTACACTGACATGACATCCACCCAT 398
QY 181 AACATGGGTTCAGATCAAGAAACCAACCAAGAGCTGAGAACTGTGTAGTCCAGGG 240
Db 399 AACATGGGACAGATCAAGAAACCAACCAAGACATGAGCTGAGAACTGTGTAGGCGCCAGG 458
QY 241 TCAGGCAAAACCCCTGACCTCATGTTTATGGCCATGCTAGCTGTAATATCCTGTGCGAGT 300
Db 459 TCAGGC---AACCCCTGACTCCATGTTCTTGGCCATGCTAGCTAGTGTCTGTGCGGT 515
QY 301 ATGATTTTCTGTCAGCAAGCAAAACATATTTGGGCATATTTCTTAACCCACCGTAGT 360
Db 516 ATGTTTCCCTGTGTAGAGG---ACATATTTGGGCATATTTCTTAATCACTTGTAGT 571
QY 361 GTGA-TCATACTCTGAAGCAGCAGCTCCTCTGAGATATATCATCATCAAGGAGCATCAGT 419
Db 572 ACAGTGGTACTCTGTAGCGACACTTCTCTGAGATATCATATGATCATGTAGCGGTGGC 631
QY 420 ACCAGGACCTCT 431
Db 632 ACCAGGATCCT 643
RESULT 10
AG183922
LOCUS 743 bp DNA linear GSS 09-JAN-2002

DEFINITION Pan troglodytes DNA, clone: RP43-057I20.TJ, genomic survey
sequence.
ACCESSION AG183922
VERSION AG183922.1 GI:16713602
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
Male BAC Library clone:RP43-057I20.TJ.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 743)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
1. .743
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-057I20.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT 231 a 148 c 174 g 167 t 23 others
ORIGIN
Query Match 24.8%; Score 248.2; DB 17; Length 743;
Best Local Similarity 81.0%; Pred. No. 4.6e-66;
Matches 289; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 1 GGAGATGATAAACCCTGTGAGTGCCCTCAAGTTGTGTGGACCATGGAATGGGAGACTG 60
Db 377 GGAGATGGACAACCATGTGGTGCCCTCAAGTTATGTGCAATCATGGAACGAGAGACTG 436
QY 61 GAGGATACATGATGCCAATCAAGCCGACGCTCCAGATGAGCCATGAGCCAGTT 120
Db 437 GAGGACCTATGGACCCCAACTATGGTGTCTCCAGTATGAGCCATGAGCCAGTT 496
QY 121 GAATCTCAATGTGAAGATGGAATGAACCGCAGAGAGTACACTGACGTCACACCTCAT 180
Db 497 GAACCTGAATGCAAGAGACAGAGCAAGGACGACGAGGTCACAGACATCAACCCCAT 556
QY 181 AACATGGGTTCAGATCAAGAAACCAACCAAGAGCTGAGAACTGTGTAGTCCAGGG 240
Db 557 AACATGAGACAGATCAATAAACCAACCAAGAGCTGAGAACTGTGTAGGACATCAGGA 616
QY 241 TCAGGCAAAACCCCTGACTCCATGTTTATGGCCATGCTAGCTGTAATATCCTGTGCGAGT 300
Db 617 TCAGGCAAAACCCCTGACTGATTTTATGGCCATGCTAGCTATAATATCATGTGCGAGT 676
QY 301 ATGATTTTCTGTCCAGCAAGCAAAACATATTTGGGCATATTTCTTAACCCACCGGT 357
Db 677 ATGTTTCCCTGTGTAGAGGCAAGACATATTTGGGCATATTTCTTAACACCTCCAGT 733

```

RESULT 11
LOCUS      BG777821
DEFINITION 60266492f1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4804708 5',
            mRNA sequence.
ACCESSION  BG777821
VERSION     BG777821.1 GI:14048138
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 574)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1651 row: j column: 05
            High quality sequence stop: 565.
FEATURES             Location/Qualifiers
     source           1..574
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4804708"
                     /clone_lib="NIH_MGC_60"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggccattatggcc
                     ); Double stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGCCGAGCGCGGCACATG-DT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T). Average
                     insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT  164 a 130 c 165 g 115 t
ORIGIN
Query Match      23.7%; Score 237.2; DB 12; Length 574;
Best Local Similarity 75.8%; Pred. No. 1.1e-62;
Matches 332; Conservative 0; Mismatches 103; Indels 3; Gaps 3;

Qy 1 GGAGATGATAAACCGTGTAGTGCCTCAAGTTGTGTGGCAGCATGGAATGGGAGCTG 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 GCAGATGATAAACCATGTGGTGCCTCAAGTTCGCGTCAACCATGGAATGGGAGACTG 192

Qy 61 GAGGATACATGGATCCCACTACAGGCCAGCTCCCTCAGTATGAGCCATGAGCCAGTT 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 GAGGAACCATGGTGGGCAACCATGGGCGCGTCCCTCTGTGTGGGCGATGAGCCAGCT 252

Qy 121 GAATCTGAATGTGAAGATGGAATGAAGACCGACGAGAGCTACACTGAGCTCAACCCCTCAT 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 GAGCCTGAGTGCATAAATGGACAGAGCGGCACTGGAGTCATGATGACATCAACCCCTAT 312

Qy 181 AACATGGGTCAGATCAAGAAACACACACAGAGCTGAGAACTGGTGTAGTGCCAGGG 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 AACCTGGGGGACACTCGAGAAACACACACAGAGCTGAGAACTGAGAGCGTCAAGG 372

Qy 241 TCAGGCAAAACCCCTGACTCCATGTTTATGCGCATGCTAGCTGTAATATCTGTGCAGT 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 373 CCA-GAAAAAACCCCTGATTCATGTTCTTGGCCATGTTAGCTATAATATGCCCCTGTGTGGT 431

```

```

Qy 301 ATGATTTTCTGTGCAGAGCAAAACATATTTGGGCATATTTTCTTCAACCCCGGTAGT 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 A-GTTTTTCCCTTCAGAGTCAAAACATATTTGGGCATAGTTCCCAATCTCCAGCAGT 490

Qy 361 GTGATC-ATACTCTGAGCAGCAGCTCTCTCTGAGATATATCATCATCAGGAGCATCAGT 419
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 491 ACAACCTGTATTTGGAGTGACACTCTCTCTGAGATATCTCTGATCATCAGGAGCATGGC 550

Qy 420 ACCAGGACCTCTAACTCC 437
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 551 TCCAGGACCCCTATGTCC 568

RESULT 12
LOCUS      AQ900691
DEFINITION HS_5010_A2_C04_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=8778 Col=8 Row=E, DNA sequence.
ACCESSION  AQ900691
VERSION     AQ900691.1 GI:6356881
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 636)
            Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
            Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
            Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
            93380589
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.bufo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.bufo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 8778 row: E column: 8
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source           1..636
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=8778 Col=8 Row=E"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT  211 a 128 c 160 g 137 t
ORIGIN
Query Match      22.4%; Score 223.8; DB 17; Length 636;
Best Local Similarity 78.8%; Pred. No. 1.8e-58;
Matches 267; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 1 GGAGATGATAAACCGTGTAGTGCCTCAAGTTGTGTGGCAGCATGGAATGGGAGCTG 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 298 GGAGATGATAAACCTCTGTGGTGTCTGTCAAGTGTGTGGCAGCAAGGAATGGGAGACTA 357

```

```

Qy 61 GAGGGATACATGGATCCCACTACAGGCCCGCTCTCCAGTATGAGCCATGAGCCAGTT 120
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 358 GAGGAACCCAGGGTGGCCCAACCATGGGCCAGTCCCTCCAGTACAAGCCATGAGCCAGCT 417
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 121 GAATCTGAATGTGAAGATGAAGACCGACGAGAGTCACTGAGGTCAACCCCTCAT 180
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 418 GAGCTTGAGTCCAAGACAGAGAACCAACCAAGTCAATGACATCAACCCCAT 477
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 181 AACATGGGGTCCAGATCAAGAAAACACACAGAAAGTCCAGAACTGGGTAGTGCCAGGG 240
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 478 AACCTGGGACAACTCAAGAAAACACGACGAAAGTCCAGAACTCTGGAGCATCAAGG 537
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 241 TCAGGCAAAAACCCCTGACTCCATGTTTATGGCCATGCTAGCTGTGAATCATCTGTGCAGT 300
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 538 CCAGACAAAACCCCTGATTCATATTCATGCTATGCTATGCTATGCTATGCTGTCAGT 597
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 301 ATGATTTTCTGTGCAGNAGCAAAAACATATTTGGGCATA 339
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 598 ATGTTTCTCTGTGCAGAGGCAAAAACATATTTGGGCATA 636
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 13
LOCUS      A1830956
DEFINITION wj80d01.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2409121 3',
            mRNA sequence.
ACCESSION  A1830956
VERSION    A1830956.1 GI:5451627
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 418)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-rc@mail.nih.gov
          Life Technologies catalog #: 11547-015
          DNA Sequencing by: Washington University Genome Sequencing Center
          Clone Distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html
          Insert Length: 578 Std Error: 0.00
          Seq primer: -40UP from Gibco.
          Location/Qualifiers
            1..418
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:2409121"
              /clone_lib="NCI-CGAP_Lym12"
              /tissue_type="lymphoma, follicular mixed small and large
              cell"
              /lab_host="DH10B"
              /note="Organ: lymph node; Vector: pCMV-SPORT6; Site:1:
              SalI; Site:2: NotI; Cloned unidirectionally. Primer:
              Oligo df. Average insert size 1.25 kb. Life Technologies
              catalog #: 11547-015"
              72 a 116 c 105 g 125 t

BASE COUNT 72 a 116 c 105 g 125 t
ORIGIN

Query Match 21.8%; Score 217.8; DB 9; Length 418;
Best Local Similarity 78.4%; Pred. No. 1.1e-56;
Matches 261; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 1 GGAGATGGATAAACCGTGTGAGTGCCTCAAGTTGTGTGGCACCATGGAATGGGAGACTG 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 368 GGAGATGACAAACCGTGTGGTGGCTCCCTCCAGTGTGTGTGACCATGGACGGGAGCCG 309
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 61 GAGGGATACATGGATCCCACTACAGGCCCGCTCTCCAGTATGAGCCATGAGCCAGTT 120

```

```

Db 308 GAGGGATCCATGGTATTCAACCGTGGGCTGTACCTCCAGTACGAGCCATGAGCCAGCG 249
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 121 GAATCTGAATGTGAAGATGAAGAACCGAGAGGTCACTGAGTCAACCTCAT 180
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 248 GAATCTGAATGCAAGACAGAACAGGGCCGACCGAGGTCACAATGACATCAACCCCAT 189
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 181 AACATGGGGTCCAGATCAAGAAAACACACAGAAAGTCCAGAACTGGGTAGTGCCAGGG 240
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 188 AACATGGGACAGATCAAGAAAACGACACAAGAGCTCAGAACTACTGGAGCGCCAGGG 129
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 241 TCAGGCAAAAACCCCTGACTCCATGTTTATGGCCATGCTAGCTGTGAATCATCTGTGCAGT 300
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 128 TCAGGCAAAAACCCCTGACTCCATGTTTATGGCCATGCTAGCTGTAGTGTCTGTGCGTC 69
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 301 ATGATTTTCTGTGCAGAGCAAAAACATATTG 333
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 68 TATAGGATCGGGTGAACCAACCGGCAATTG 36
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 14
LOCUS      BF725513
DEFINITION bxl6g05.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo
            sapiens cDNA clone bxl6g05 5', mRNA sequence.
ACCESSION  BF725513
VERSION    BF725513.1 GI:12041424
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 524)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL   Wistow, G.J., Bernstein, S., Behal, A. and Smith, D.
COMMENT   NEIBANK: EST analysis and bioinformatics for ocular genomics
          Invest. Ophthalmol. Vis. Sci. 41, (2000) In press
          Contact: Wistow G
          Section on Molecular Structure and Function
          National Eye Institute
          6/331, NIH, Bethesda, MD 20892-2740, USA
          Tel: 301 402 3452
          Fax: 301 496 0078
          Email: graeme@helix.nih.gov
          Plate: 16 row: g column: 05
          Seq primer: M13Rpl reverse primer (ABI).
          Location/Qualifiers
            1..524
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="bxl6g05"
              /clone_lib="Human Iris cDNA (Un-normalized, unamplified):
              BX"
              /tissue_type="Iris"
              /dev_stage="Adult"
              /lab_host="EMDH10B"
              /note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris
              tissue was pooled from 10 individuals ranging in age from
              4-80 years and RNA was extracted. From this pooled sample
              an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A
              directionally cloned cDNA library in the pCMVSPORT6 vector
              was constructed at Life Technologies, essentially
              following the protocols of the SuperScript plasmid system
              full details of which are contained in the manufacturer's
              instruction manual (http://www.lifetech.com/). First
              strand synthesis was carried out using a Not I
              primer-adaptor [5'-pCACTAGTCTAGATCGGAGCGGCCCT(T)15-3'
              ]. Not I/plunt end inserts were cloned into the Not I/EcoR
              V sites in the vector. EST analysis was performed on the
              unamplified library at the NIH Intramural Sequencing
              Center (NISC)."
              148 a 133 c 160 g 83 t

BASE COUNT 148 a 133 c 160 g 83 t
ORIGIN

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:29:47 ; Search time 1108.8 seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724A-3_COPY_99000_100000
Perfect score: 1001
Sequence: 1 tggcgcgcgtgtagctcc.....ttttatccaccatcaactaa 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	303.2	30.3	648	12 BE887242	BE887242 601508554
2	283.2	28.3	763	13 BI551044	BI551044 603195787
3	212.6	21.2	559	12 BG722262	BG722262 602693433
4	206	20.6	526	12 BF433607	BF433607 7q59g03.x
5	206	20.6	662	12 BE948806	BE948806 UI-M-BH3-
6	202	20.2	1143	11 AK006906	AK006906 Mus muscu

c	7	195.4	19.5	517	10	AW293335
	8	192	19.2	470	9	AI982914
	9	178	17.8	934	14	BM800032
	10	165.6	16.5	845	13	BI561035
	11	163.8	16.4	372	14	U46290
	12	156.6	15.6	757	13	BI086746
c	13	136	13.6	458	10	BE670624
	14	132.2	13.2	678	13	BI462516
	15	118.8	11.8	337	14	T07679
	16	118.6	11.8	487	9	AA991438
c	17	117.4	11.7	636	17	AG066329
	18	116.6	11.6	600	12	BF678990
c	19	116.4	11.6	480	17	AQ122880
c	20	116.2	11.6	537	17	AQ390912
c	21	115.6	11.5	693	17	AQ315247
	22	114.4	11.4	206	13	BM456815
	23	114.4	11.4	641	12	BF674369
c	24	113.8	11.4	598	9	AL134441
c	25	113.4	11.3	490	10	AW265073
c	26	113.2	11.3	651	17	BE6038
c	27	113.2	11.3	2169	11	AF305815
c	28	113	11.3	737	9	AL602864
c	29	112.8	11.3	435	17	AQ132070
c	30	112.8	11.3	571	17	B95510
c	31	112.8	11.3	651	17	AG041999
c	32	112.6	11.2	301	9	AI311276
c	33	112.4	11.2	657	17	B02497
c	34	112.2	11.2	541	17	AQ456714
c	35	112.2	11.2	600	17	AQ393075
c	36	112.2	11.2	606	17	AQ345963
c	37	112.2	11.2	2580	11	BC024736
c	38	112	11.2	570	9	AL703987
c	39	112	11.2	677	12	BG612299
c	40	111.8	11.2	460	17	AQ076891
c	41	111.8	11.2	673	17	AQ274719
c	42	111.4	11.1	515	9	AA665160
c	43	111.2	11.1	358	17	B72389
c	44	111.2	11.1	535	17	AQ391712
c	45	111.2	11.1	628	17	AG182392

ALIGNMENTS

RESULT 1
BE887242
LOCUS BE887242 601508554F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3910044 5', linear EST 20-OCT-2000
DEFINITION mRNA sequence.
ACCESSION BE887242
VERSION BE887242.1 GI:10342334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 648)
AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9724 row: 1 column: 13
High quality sequence stop: 588.
Location/Qualifiers 1. 648

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3910044"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."
Average insert size 2.1 kb.
BASE COUNT      189 a 141 c 138 g 180 t
ORIGIN

Query Match      30.3%; Score 303.2; DB 12; Length 648;
Best Local Similarity 79.5%; Pred. No. 5.2e-49;
Matches 420; Conservative 0; Mismatches 78; Indels 30; Gaps 4;

QY 502 AGGAGCTGAGCGGGGATCATCTCCAGGAAGGGGCTGAGTAGGACTGCGAGCGCG---558
|||||
Db 1 AGGGAGCTGAGCGGGGAGGACCTCCCGGGGACCCAGCGAAGGAGGAGCGCGCGCGCGC 60
|||||
QY 559 -----CACATCCTCTCTTTACCCGGGGATGTGCAGGATTACCGTGAATCATGACTC 610
|||||
Db 61 CCCAGCATCAGCTCTCTATTTCCCTGGGGATGTGCAGATTACCATGAATATGACTC 120
|||||
QY 611 GTATCTCTCGGAATTACCAATGGGAAATTTGGAGTCTAGAAATTAFTGCTCCATTTAG 670
|||||
Db 121 GTCATCTGAGAATTATCAATGGGAAACTGGAGTCTAGAAAATGTTGCTACCATTTAG 180
|||||
QY 671 CCCACGGTTCCCACTAGCTGTATTGGGGTGATGAAGTGTCCCGAA-----C 719
|||||
Db 181 CCCACGGTTCCCACTAGCTGTATTGGGTGATCAAAATGTTCCCGAATGCAATTTGCACA 240
|||||
QY 720 AAATGGCTGCCATGATAGTTTCTGAAAAGTAACATGTTGGTTTCCAGAACACAATA 779
|||||
Db 241 AATTACGCTGCTATGACAAATTTGGAAGTAACATGTTGGTCCCGCCAGACACAATA 300
|||||
QY 780 CAGACTCTGGAGCTTTAAGCACCTTTATATATTAFTAGTTAA-----TGCTTTTAAGT 833
|||||
Db 301 CTGACTTTGGAGCTTTAAGCACCTTTATATGTTATTACGTTTAATGCTTTTAATTAAGT 360
|||||
QY 834 CAGAGTAGTTTATCAAGGAAATTTGAATGATTGGAATGAAGGACTCCACAGCACTAAT 893
|||||
Db 361 CAGAATAGTTTATCAAGGAAAGTTTGAATGTTTGAATGAAGGACTCCATAGCATCTAAC 420
|||||
QY 894 TGTAGATGTCCTTCTCTATACATACAAATCATTTCCAGGAAGGAAAGATAGGACCT 953
|||||
Db 421 TGTAGA--TCCAGCTCTCTCATACTACGATGCTTCCAGGAGGAAAGCAAGTCAGGACCT 478
|||||
QY 954 TTGAAAATCTGATGATCGGCGCATGTGTTTATCCACCATCACTAA 1001
|||||
Db 479 GTGACAACTCTGATGAGTGGCATGAGTTTATCCACCATCACTAA 526
|||||

RESULT 2
BI551044
LOCUS
DEFINITION
603195787F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5275580 5',
mRNA sequence.
BI551044
ACCESSION
BI551044.1 GI:15438356
VERSION
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 763)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
```

```
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11695 row: m column: 21
High quality sequence stop: 692.
FEATURES
Location/Qualifiers
source
1..763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5275580"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      236 a 143 c 148 g 236 t
ORIGIN

Query Match      28.3%; Score 283.2; DB 13; Length 763;
Best Local Similarity 83.6%; Pred. No. 3.6e-45;
Matches 363; Conservative 0; Mismatches 53; Indels 18; Gaps 3;

QY 584 TCCAGGATTACCGTGAATCATGCTCGTCATCTCCGCAATTACCAATGGGAAATTTGA 643
|||||
Db 37 TTCAGAAATTACCATGAATATGACTCGTCATCTCGAGAAATTATCAATGGGAAATCGA 96
|||||
QY 644 GTCTAGAAATTAATGCCCTCCATTTTAGCCCCACCGGTTCCCGAGTAGCTGTATTTGGGGTGA 703
|||||
Db 97 GTCTAGAAATGTGTGCTACCATTTTAGCCCCACCGGTTCCCGAGTAGCTGTATTTGGGGTGA 156
|||||
QY 704 TGAAGTGTCCCGAA-----CAATGGCTGCCATGATAGTTTCTCGAAAGTA 752
|||||
Db 157 TAAATGTGTCCCGAATGCAATTTGCACAAATTCAGCTGCTATGACAAATTTGTGAAAGTA 216
|||||
QY 753 ACATGTTTGGTTTCCAGAACACATACAGACTCTGGAGCTTTTAAAGCACCTTTATATGT 812
|||||
Db 217 ACATGTTTGGTCCCGCAGACACATAGCTTTGGAGCTTTTAAAGCACCTTTATATGT 276
|||||
QY 813 TATTAGTTAATGCT-----TTTAAAGTCAGAGTAGTTTATCAAGGAAATTTGAATGATT 867
|||||
Db 277 TATTAGTTAATGCTTTTAAATTTAAGTCAGATAGTTTATCAAGGAAATTTGAATGTTT 336
|||||
QY 868 GGAATAAGGACTCCAGCAGCATTAATTTAGATGTCCTCAATTTCTTCATCACTACAATCA 927
|||||
Db 337 GGAATAAGGACTCCATAGCATCTAATCTAGTA--TCCAGTCTCTTCATCACTACGAAATG 394
|||||
QY 928 TTTCCAGGAAGGAAAGATAGGACCTTTCAAAATCTGATGATGCGCCATGTGTTTTTA 987
|||||
Db 395 TTGCAGGAGGAGAAAGTAGGAGGACCTGTGAAAAATCTGATGAGTCGCGCATGAGTTTTTA 454
|||||
QY 988 TCCACCATCACTAA 1001
|||||
Db 455 TCCACCATCACTAA 468

RESULT 3
BG722262
LOCUS
DEFINITION
602693433F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4825616 5',
mRNA sequence.
BG722262
ACCESSION
BG722262.1 GI:14001449
VERSION
EST.
KEYWORDS
```

```

SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 559)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgraphs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10739 row: a column: 09
High quality sequence stop: 559.
FEATURES
Location/Qualifiers
1..559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4825616"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 93 a 200 c 173 g 93 t
ORIGIN
Query Match 21.2%; Score 212.6; DB 12; Length 559;
Best Local Similarity 75.7%; Pred. No. 1.8e-31;
Matches 308; Conservative 0; Mismatches 84; Indels 15; Gaps 3;
QY 284 GGTGTTGCGGCGCGGTGTTGCTCTGCTGCCAGAGTTTGGAACTGGAGATGCGCTC 343
D 153 GCTCTGCGGTCGCGCGGTGTTGCTCTGCTGCCAGAGTCGGAATCGGAACGCTC 212
QY 344 TTCCTTCTCAGCAGACAGACCATG-AGCCTAGCGGAGCGCGCGGTTCGCGAAGCTCCCC 402
D 213 GTCTGCTCTCAGGCGAGAACCATGAACCGCGGCGGCGCGCGCGGCGGATGCTCCT 272
QY 403 CTCCGCCAACGCGCGCTCTCAGAGCGGT--CCGTGCGCGCGCTGCGGAGCGTACC 459
D 273 CTCCGCGAGTGGCGCGCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 332
QY 460 CGCAGCGCAGTGCAGAGTGTCTCTGTTGGCGGTGACAGGAGGAGGAGTGGAGCGGCGGA 519
D 333 CGCAGCGCAGCAGCAATGTTGCTCTGTTGGCGGCGCGGCGGAGGAGTGGAGCGGCGAG 392
QY 520 TCATCTCCAGGAAGCGGCTGAGTAGAAGTGCAGCGCG-----CACATCTCTCT 568
D 393 ACCTCCCGGGGACCCGAGGAGGAGGCGCGGAGGCGCGCGCGCGCGCGCGCGCTCT 452
QY 569 GTTTTACCGGGGAGTGCAGGATACCTGGAATCATGACTCGTTCCTCTCGAATTACC 628
D 453 ATTTCCCTGGGGATGTGCAGAAATTTACCATGAATTTATGACTCGTATCTCTGAGAATATC 512
QY 629 AATGGGAAATTTGGAGTCTAGAAATTTATGCTCTCATTTTACGCCAC 675
D 513 AATGGGAAACTGGAGTCTAGAAATTTGCTTACCATTTGAGCCAC 559
RESULT 4
BASE COUNT 161 a 96 c 77 g 192 t
ORIGIN
Query Match 20.6%; Score 206; DB 12; Length 526;
Best Local Similarity 100.0%; Pred. No. 3.6e-30;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 796 TAAGCACCTTTATATGTTATTTAGTTAATGCTTTTAAGTCAGAGTAGTTTATCAAGGAAA 855
D 526 TAAGCACCTTTATATGTTATTTAGTTAATGCTTTTAAGTCAGAGTAGTTTATCAAGGAAA 467
QY 856 ATTTGAATGATTGGAATAAGGACTCCACAGCATCTAATTTAGATGTCCTAATTTCTTCA 915
D 466 ATTTGAATGATTGGAATAAGGACTCCACAGCATCTAATTTAGATGTCCTAATTTCTTCA 407
QY 916 TACTACAAATCATTTCCAGGAAGGAAAGATAGGACCTTTTGAAAATATGTATGATCGGC 975
D 406 TACTACAAATCATTTCCAGGAAGGAAAGATAGGACCTTTTGAAAATATGTATGATCGGC 347
QY 976 CATGTGTTTTTATCCACCATCACTAA 1001
D 346 CATGTGTTTTTATCCACCATCACTAA 321
RESULT 5
BE948806
LOCUS
DEFINITION UI-M-BH3-avj-h-01-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone
EST 03-OCT-2000
linear mRNA
662 bp
NIH_BMAP_M_S4
```

UI-M-BH3-avj-h-01-0-UI 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE948806
BE948806.1 GI:10526565
EST.

house mouse.
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 662)
Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL
MEDLINE
COMMENT

Genome Res. 6 (9), 791-806 (1996)
9704477

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: meste@mail.nih.gov

Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers

1..662

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH3-avj-h-01-0-UI"

/dev_stage="27-32 days"

/lab_host="DHI0B (Life Technologies)"

/note="(Vector: p7T30-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI: The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DHI0B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)

TAG_SEQ=None found"

BASE COUNT 163 a 168 c 166 g 164 t 1 others

ORIGIN

Query Match 20.6%; Score 206; DB 12; Length 662;
Best Local Similarity 71.1%; Pred No. 3.3e-30;
Matches 323; Conservative 0; Mismatches 110; Indels 21; Gaps 3;

Qy 554 GCCGCCACATCCTCTCTTTACCCGGGGATGTCAGGATTACCGTGAATCATGACTCGTC 613
Db 192 GGCACACGCTCTCTACTTCCCTGGGACGTGCAGAATTACCATGAGACTACTCGTC 251
Qy 614 ATCTCGGAATTACCAATGGGAATTTGGAGTCTAGAATATTGCCCTCCATTTAGCCC 673
Db 252 ATCTCGGAATTACCAATGGGAATTTGGAGTCTAGAATATTGCCCTCCATTTAGCCC 311
Qy 674 ACCGTTTCCCAAGTGTATTTGGGTTGATGAAGTGTCCCGAA-----CAAA 722
Db 312 GCGTTTCCCAAGTGTATTTGGGTTGATGAAGTGTCCCGAAATGATTTGCACAAT 371
Qy 723 TCGCTGCCATGATAGTTTCTGAAAAGTACATGTTGGTTTCCCAAGACATATACAG 782
Db 372 TCAGCTGCTATGACAAATTTTGTGAAAAGACATGTTTGGTGCACCAAGACATCTCCG 431
Qy 783 ACTCTGGAGCTTTTAAGCACCTTTATATGTTATTTAGTTATTAAGTCTTTTA-----AGTCAGA 837
Db 432 ACTTTGGAGCTTTTAAGCACCTTTATATGTTATTTAGTTATTAAGTCTTTTAACCTGACTCAA 491
Qy 838 GTAGTTTATCAAGGAAAATTTGAATGATTTGAATAAGGACTCCACAGCATCTAATTGTA 897
Db 492 ACGGTATCTGTTCAAGAACAGGAGCTGTTGGAATAAGGATTTGAAGCATCTAAGTGTG 551
Qy 898 GATGTCCTCAATTTCTCTACTACTACAAATCATTTCCAGGAAGAAAGATAGGACCTTTGA 957
Db 552 AGT-----CTAATCTCTTCTACTAGTAATGGTGGCCAAAAGGAAAACGAGAGACCTGTGA 606
Qy 958 AAAATCTCATGATCGGCGCATGTGTTTATCCA 991
Db 607 GCATGTTGATGAGCCTTCCATGAGTTTCTCTCCA 640
RESULT 6
LOCUS AK006906 1143 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700066M21:related to CG7687 PROTEIN, full insert
sequence.
ACCESSION AK006906
VERSION AK006906.1 GI:12840206
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
clone:1700066M21.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Tanaka,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 517)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..517
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2726801"
 /clone_lib="NCI-CGAP_Sub4"
 /lab_host="DH10B (Life Technologies)"
 /note="vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NCI-CGAP Sub4 library is a subtracted library derived from the NCI-CGAP_Sub2 library which is a subtracted library derived from the NCI-CGAP_Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI-CGAP libraries: NCI-CGAP_Co4, NCI-CGAP_Pr22, NCI-CGAP_Pr28, NCI-CGAP_Co10, NCI-CGAP_Co16, NCI-CGAP_Kid5, NCI-CGAP_Kid12, NCI-CGAP_Kid3, NCI-CGAP_Kid11, NCI-CGAP_Lym2, NCI-CGAP_Br2, NCI-CGAP_Co8, NCI-CGAP_CLL1, NCI-CGAP_Le12, NCI-CGAP_Brn23, NCI-CGAP_Lu5, NCI-CGAP_Lu24, NCI-CGAP_Lu19, NCI-CGAP_GC4, NCI-CGAP_GC6, NCI-CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
 NCI-CGAP_Kid3 pool 1 : LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775, 1500552-1502855) NCI-CGAP_Kid5 pool 1 : LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831, 1471368-1472903, 1492104-1493255) NCI-CGAP_Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1520904-1522439) NCI-CGAP_GC4 pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743) NCI-CGAP_Pr22 pool 1 : LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615) NCI-CGAP_Co10 pool 1 : LLAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351) Subtraction was performed as previously described [Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.]
 TAG-LIB=NCI-CGAP_Lu5
 TAG_TISSUE=lung
 TAG_SEQ=CAAC*

BASE COUNT 159 a 75 g 188 t

ORIGIN

Query Match 19.5%; Score 195.4; DB 10; Length 517;
 Best Local Similarity 99.5%; Pred. No. 4e-28;
 Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 805 TTATATGTTATTAGTCTTTTAACTCAGAGTAGTTTATCAAGGAAATTTGAATG 864
 Db 517 TTATATGTTATTAGTCTTTTAACTCAGAGTAGTTTATCAAGGAAATTTGAATG 458

QY 865 ATTCGAATAAGGACTCCACAGCATCTAATTGTAGATGTCCAATTCTTCTCATACTACAAA 924
 Db 457 ATTCGAATAAGGACTCCACAGCATCTAATTGTAGATGTCCAATTCTTCTCATACTACAAA 398
 QY 925 TCATTTCCAGGAAGAAAGATAGGACCTTTGAAAAATCTGATGATCGGCGCATGTGTTT 984
 Db 397 TCATTTCCAGGAAGAAAGATAGGACCTTTGAAAAATCTGATGATCGGCGCATGTGTTT 338
 QY 985 TTATCCACCACATCACTAA 1001
 Db 337 TTATCCACCACATCACTAA 321
 RESULT 8
 AI982914
 LOCUS w229f09.x1 NCI-CGAP_Brn53 Homo sapiens cDNA clone IMAGE:2559497 3', mRNA linear EST 09-MAR-2000
 DEFINITION mRNA sequence.
 ACCESSION AI982914
 VERSION AI982914.1 GI:5810133
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 470)
 NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index
 Unpublished (1998)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert length: 3011 Std Error: 0.00
 Seq primer: -40UP from Glibco
 High quality sequence stop: 408.
 FEATURES
 Location/Qualifiers
 1..470
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2559497"
 /clone_lib="NCI-CGAP_Brn53"
 /tissue_type="three pooled meningiomas"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."
 BASE COUNT 79 a 167 c 142 g 80 t 2 others
 ORIGIN
 Query Match 19.2%; Score 192; DB 9; Length 470;
 Best Local Similarity 74.7%; Pred. No. 1.9e-27;
 Matches 284; Conservative 0; Mismatches 82; Indels 14; Gaps 3;
 QY 284 GGTTTTGGGTGCGCGGTGTTGCTGCTGCTGCCACAGTTGGAACTGGAGATGCCTC 343
 Db 91 GCTCCTGGGTGCGCGGTGTTGCTGCTGCTGCCAGCTCGGAATCGGAACGCCCTC 150
 QY 344 TTCCTTCTCTCAGGACAGAACCATG-AGCCTAGCGGAGCGCGGTTGCGGAAGTCCCC 402
 Db 151 GTCCTGCTCTCAGGCGAGAACCATGAACCCGGCGGAGCGCGCGCATGCTCCT 210
 QY 403 CTCGCCCAACGGCGCTCTCCTCAGAGCGTCC--GTGCCCGCGCTGCCGGAGCTGACCC 460
 Db 211 CTCGCCGAGGTGCGCGCGGCTGCTGCTGCTGCTTCCACCGGTGCTGGAGCGGATCC 270

Qy	461	GCAGC	-GCAGTGCAGAGTTGCTCTCTGTTGGCGGTGACACAGGAGGGAGTGCAGCGCGGGA	519
Db	358	GCAGCTGCAGGAAGAAATTGCTCTGTTGGCGGCGCGGGAGGAGTGCAGCGCGCAGG	417	
Qy	520	TCATCTCAGGAGCGGGCTGA	--GTAGGAACTGCAGCCG	566
Db	418	ACCTCCCGGGAGCCACAGCAGAGAGGAGGAGCGACGCCGCCGCCAGCATCAGCTCCT	477	
Qy	567	CTCTTTACCCGGGATGTGCAGGATTACCGTGAATCATGACTCGTCATCTCGGAATTA	626	
Db	478	CTATTTCCCTGGGATGTGCAGAAATACCATGAATATGACTCGTCATCTCGAATAA	537	
Qy	627	CCAAATGGG-AAATTTGAGTCTAGAA-ATTATTCGCCCTCCATTTAGCC	672	
Db	538	TCAATGGGAAACATGGAGTCTAGAACAAATGTTGCTACCATTTTAGAC	585	
RESULT 11				
U46290				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
COMMENT				
FEATURES				
source				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 257; Conservative				
Qy 284				
Db 1				
Qy 344				
Db 61				
Qy 403				
Db 121				
Qy 461				
Db 181				
Qy 520				
Db 240				

Qy	521	CATCTCCAGGAAGGGGCTGAGTAGGAACTGACGCCG-----CACATCCTCTTC	569
Db	241	CCTCCCGGGGACCCAGGAGGAGGAGCGGAGNCGCCGCCCCAGNATCACGTCCTCTA	300
Qy	570	TTTATCCCGGGGATGTGAGGATTACCTGAAATCATCATCTCGTCATCTCGGAATTACCA	629
Db	301	TTTCTCTGGGATGTGNAGAAATTACCATGAATTTATGACTCGNCACTCGAGAATTATCA	360
Qy	630	ATGGGAAAATTG 641	
Db	361	ATGGGAAAACGTG 372	
RESULT 12			
BI086746			
LOCUS			
DEFINITION	602850217F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4991524 5', mRNA sequence.	757 bp	linear EST 20-JUN-2001
ACCESSION	BI086746		
VERSION	BI086746.1	GI:14505076	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 757)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM1009 row: b column: 05 High quality sequence start: 7 High quality sequence stop: 633. Location/Qualifiers		
FEATURES			
source	1. 757 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4991524" /clone_lib="NIH_MGC_10" /cell_line="MGC36" /lab_host="DH10B" /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."		
BASE COUNT	239 a 134 c 166 g 218 t		
ORIGIN			
Query Match	15.6%	Score 156.6; DB 13; Length 757;	
Best Local Similarity	84.9%	Pred. No. 1.le-20;	
Matches	213; Conservative	0; Mismatches 29; Indels 9; Gaps 3;	
Qy	757	GTTTGGTTTCCAGACACATACGACTCTGGAGCTTTTAAGCACCTTTATATGTTATT	816
Db	1	GTTAGGTGCCCCAGAACAACTATGACT-TGGAGCTTTGAAGCACCTTTATATGTTATT	59
Qy	817	AGTTAATGCTTTTAAGTCA-----GAGTAGCTTTATCAAGGAAAAATTGAATGATTGGA	870
Db	60	AGTTAATGCTTTGAATTTAAGTCAGAACTACTTTATCAAGAAAAAGTTGAATGTTTGA	119
Qy	871	ATAAGGACTCCACAGCATCTAATTTGTAGATGCCAATTTCTCTCATACTACAAATCATTT	930
Db	120	ATAAGGACTCCATAGCATCTAAGCTAGTA--TCCAGTCTTCTCTACTACTACGAATGGTTG	177

Qy	931	CCAGGAAGGAAAAAGATAGGACCTTTGAAAAATCTGATGGATCGGCCCATGTGTTTTATPCC	990
Db	178		
		CCAGGGAGAAAAAGTGAGGACCTGTGAAAAATCTGATGAGTCTGCCATGAGTTTTATPCC	237
Qy	991	ACCATCACTAA 1001	
Db	238		
		ACCATCACTAA 248	
RESULT 13			
BE670624/c			
LOCUS	BE670624	783708.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3284655 3', mRNA 458 bp linear EST 08-SEP-2000	
DEFINITION		mRNA sequence.	
ACCESSION	BE670624		
VERSION	BE670624.1	GI:10031165	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 458)	
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL		Unpublished (1997)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -400P from Gibco.	
FEATURES			
source		1. .458 Location/Qualifiers	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:3284655"	
		/clone_lib="NCI_CGAP_Lu24"	
		/tissue_type="carcinoid"	
		/lab_host="DH10B"	
		/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT	137 a	84 c	70 g
ORIGIN			
		Query Match	13.6%; Score 136; DB 10; Length 458;
		Best Local Similarity	100.0%; Pred. No. 1.3e-16;
		Matches 136; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	866	TTGGAATAAGGACTCCACAGCATCTAATTGTAGATGTCCAATCTTCTCATACTACAAAT	925
Db	458		
		TTGGAATAAGGACTCCACAGCATCTAATTGTAGATGTCCAATCTTCTCATACTACAAAT	399
Qy	926	CATTTCAGGAAGAAAGATAGGACCTTTGAAAAATCTGATGGATCGGCCCATGTGTTT	985
Db	398		
		CATTTCAGGAAGAAAGATAGGACCTTTGAAAAATCTGATGGATCGGCCCATGTGTTT	339
Qy	986	TATCCACCATCACTAA 1001	
Db	338		
		TATCCACCATCACTAA 323	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:29:47 ; Search time 1108.8 Seconds
(without alignments)
14620.861 Million cell updates/sec

Title: US-09-708-724a-3_COPY_10000_11000
Perfect score: 1001
Sequence: 1 caaaatttcagtttaggaaga.....gcagcacataatgatcatg 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. | No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	35.4	637	9 AL705037	AL705037 DKFZp686C
2	354	35.4	694	9 AL704947	AL704947 DKFZp686L
3	206.4	20.6	949	14 BQ230091	BQ230091 AGENCOURT
c 4	188.8	18.9	391	17 AQ093103	AQ093103 HS_3020_B
c 5	119	11.9	371	9 AA417076	AA417076 zul3b12.s
6	113.6	11.3	397	9 AA417055	AA417055 zul3b12.r

c 7	89.2	8.9	432	17 B50982	B50982 CIT978SK-19
c 8	85.6	8.6	726	17 AG153358	AG153358 Pan trogl
c 9	84.6	8.5	296	14 T39269	T39269 Ya03b03.r2
10	83.4	8.3	495	17 AQ134321	AQ134321 HS_3051_B
11	83.4	8.3	518	17 AQ169197	AQ169197 HS_3175_B
12	83.2	8.3	689	17 AQ779816	AQ779816 HS_5572_B
c 13	82.8	8.3	689	17 AG183488	AG183488 Pan trogl
14	82.4	8.2	391	14 T05554	T05554 EST03443 Fe
15	82.4	8.2	550	17 AQ468873	AQ468873 HS_5140_B
c 16	82.2	8.2	670	10 AV731805	AV731805 AV731805
17	81.8	8.2	522	17 AQ547767	AQ547767 RPCI-11-3
18	81.2	8.1	661	17 AQ127758	AQ127758 Pan trogl
19	80	8.0	455	17 AQ698289	AQ698289 HS_5551_B
c 20	79.6	8.0	750	14 BM981428	BM981428 UI-CF-EN1
c 21	79.6	8.0	514	12 BF080764	BF080764 231769 MA
22	79	7.9	514	17 AQ118584	AQ118584 HS_3012_A
c 23	78.4	7.8	626	10 AQ294157	AQ294157 UI-H-BI2
24	77.6	7.8	696	17 AG031398	AG031398 Pan trogl
c 25	77.4	7.7	614	17 AQ240524	AQ240524 CIT-HSP-2
c 26	77.2	7.7	395	9 AI263349	AI263349 qql3a08.x
27	77	7.7	524	17 AQ804711	AQ804711 HS_3094_A
c 28	76.4	7.6	490	17 AQ819320	AQ819320 HS_5295_A
c 29	76	7.6	448	17 B60070	B60070 CIT-HSP-387
c 30	75.8	7.6	330	17 B86003	B86003 RPCI11-21L1
c 31	75.8	7.6	399	9 AI203367	AI203367 qr24h10.x
c 32	75.6	7.6	729	17 AG013623	AG013623 Homo sapi
c 33	75.4	7.5	665	17 AG171538	AG171538 Pan trogl
34	75.2	7.5	391	12 BE701171	BE701171 PM3-NN012
35	75.2	7.5	470	17 B98998	B98998 CIT-HSP-228
36	75	7.5	409	17 AQ712786	AQ712786 HS_5387_A
c 37	74.8	7.5	388	9 AA581747	AA581747 nn48a05.r
c 38	74.4	7.4	729	17 AG105429	AG105429 Pan trogl
c 39	74.2	7.4	744	17 AQ480400	AQ480400 RPCI-11-2
c 40	73.8	7.4	401	14 R96184	R96184 vt84g10.r1
c 41	73.8	7.4	471	14 BQ234610	BQ234610 hda7f06.g
c 42	73.8	7.4	530	17 AQ420934	AQ420934 RPCI-11-2
c 43	73.8	7.4	536	17 AQ415616	AQ415616 RPCI-11-2
c 44	73.8	7.4	540	17 AQ785791	AQ785791 HS_5547_A
c 45	73.8	7.4	574	14 BQ234796	BQ234796 hda50b02.g

ALIGNMENTS

RESULT 1
AL705037
LOCUS AL705037 637 bp mRNA linear EST 22-MAR-2002
DEFINITION DKFZp686C1134_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
XREF DKFZp686C1134 5', mRNA sequence.
ACCESSION AL705037
VERSION AL705037.1 GI:19688392
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 637)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Weill, B. and Wiemann, S.
TITLE EST (Ottenwaelder, B., Obermaier, B., Mewes, H.W., Weill, B. and Wiemann, S.)
JOURNAL Unpublished (2001)
COMMENT Contact: Ottenwaelder B
MIPS Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZp686C1134) is available at the R2PD in Berlin.
Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin- Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES
source

1. .637
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686C1134"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: priplex2; Site_1: SfiIA; Site_2: SfiIB;
cdna-collection"

BASE COUNT 191 a 116 c 133 g 196 t 1 others
ORIGIN

Query Match 35.4%; Score 354; DB 9; Length 637;

Best Local Similarity 99.7%; Pred. No. 7.2e-64; Indels 1; Gaps 1;
Matches 365; Conservative 0; Mismatches 0;

Qy 637 TCCTGGGTATAGTGTAGTTGTAATCTGTGTTTACATACAGGATAACATAAAACAAAGG 696

Db 1 TCCTGGGTATAGTGTAGTTGTAATCTGTGTTTACATACAGGATAACATAAAACAAAGG 60

Qy 697 TAAACAATAAATAAAACAGACAGCAAACTCAACTAATAGTGTGGCATGTGTGACA 756

Db 61 TAAACAATAAATAAAACAGACAGCAAACTCAACTAATAGTGTGGCATGTGTGACA 120

Qy 757 GTGAAGACAGGAGCTCATATAAAATGGAGTGAAC-TTTTGAGCTAAATCAATGTCC 815

Db 121 GTGAAGACAGGAGCTCATATAAAATGGAGTGAAC-TTTTGAGCTAAATCAATGTCC 180

Qy 816 TGTGTGTTCTGTTCTACATCAGACTCTATAGTGCATTTCTCAGTTAGTGTGTTT 875

Db 181 TGTGTGTTCTGTTCTACATCAGACTCTATAGTGCATTTCTCAGTTAGTGTGTTT 240

Qy 876 TTATCCTTGCCTGCTAGTAAGTCCAGAGGAGATTTTCTAAACTGGTGAGGAACAGG 935

Db 241 TTATCCTTGCCTGCTAGTAAGTCCAGAGGAGATTTTCTAAACTGGTGAGGAACAGG 300

Qy 936 TAGAAGTGTAAAGTGAGACAACCTCCCTGCTCATTTGCCAAGTGGCAGCATAAATGT 995

Db 301 TAGAAGTGTAAAGTGAGACAACCTCCCTGCTCATTTGCCAAGTGGCAGCATAAATGT 360

Qy 996 ATCATG 1001

Db 361 ATCATG 366

RESULT 2

AL704947

LOCUS DKFZp686L2233_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFZp686L2233 5', mRNA sequence.

ACCESSION AL704947

VERSION AL704947.1 GI:19688302

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 694)
Ottawaelder, B., Obermaier, B., Mewes, H.W., Well, B. and
Wiemann, S.

EST (Ottawaelder, B., Obermaier, B., Mewes, H.W., Well, B. and Wiemann
, S.)

Unpublished (2001)

Contact: Ottawaelder B

MIPS

JOURNAL

COMMENT Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Medigenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No sl sequence
available.

This clone (DKFZp686L2233) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin- Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES
source

1. .694
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686L2233"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cdna-collection"

BASE COUNT 205 a 131 c 140 g 216 t 2 others
ORIGIN

Query Match 35.4%; Score 354; DB 9; Length 694;

Best Local Similarity 99.7%; Pred. No. 7e-64; Indels 1; Gaps 1;
Matches 365; Conservative 0; Mismatches 0;

Qy 637 TCCTGGGTATAGTGTAGTTGTAATCTGTGTTTACATACAGGATAACATAAAACAAAGG 696

Db 1 TCCTGGGTATAGTGTAGTTGTAATCTGTGTTTACATACAGGATAACATAAAACAAAGG 60

Qy 697 TAAACAATAAATAAAACAGACAGCAAACTCAACTAATAGTGTGGCATGTGTGACA 756

Db 61 TAAACAATAAATAAAACAGACAGCAAACTCAACTAATAGTGTGGCATGTGTGACA 120

Qy 757 GTGAAGACAGGAGCTCATATAAAATGGAGTGAAC-TTTTGAGCTAAATCAATGTCC 815

Db 121 GTGAAGACAGGAGCTCATATAAAATGGAGTGAAC-TTTTGAGCTAAATCAATGTCC 180

Qy 816 TGTGTGTTCTGTTCTACATCAGACTCTATAGTGCATTTCTCAGTTAGTGTGTTT 875

Db 181 TGTGTGTTCTGTTCTACATCAGACTCTATAGTGCATTTCTCAGTTAGTGTGTTT 240

Qy 876 TTATCCTTGCCTGCTAGTAAGTCCAGAGGAGATTTTCTAAACTGGTGAGGAACAGG 935

Db 241 TTATCCTTGCCTGCTAGTAAGTCCAGAGGAGATTTTCTAAACTGGTGAGGAACAGG 300

Qy 936 TAGAAGTGTAAAGTGAGACAACCTCCCTGCTCATTTGCCAAGTGGCAGCATAAATGT 995

Db 301 TAGAAGTGTAAAGTGAGACAACCTCCCTGCTCATTTGCCAAGTGGCAGCATAAATGT 360

Qy 996 ATCATG 1001

Db 361 ATCATG 366

RESULT 3

BO230091

LOCUS AGENCOURT_7575432 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6049217

DEFINITION 5', mRNA sequence.

ACCESSION BO230091

VERSION BO230091.1 GI:20411491

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 949)
NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13299 row: h column: 18
High quality sequence stop: 587.

FEATURES

Location/Qualifiers
1. .949
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6049217"
/clone_lib="NIH_MGC_72"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 266 a 234 c 222 g 226 t 1 others
ORIGIN

Query Match 20.6%; Score 206.4; DB 14; Length 949;
Best Local Similarity 69.1%; Pred. No. 3.7e-33;
Matches 307; Conservative 0; Mismatches 116; Indels 21; Gaps 1;

QY 294 AAAGAGAAATGAAGATACAAATGTGCACAGAGAGAAATGGCCACATGAGGACACAAAT 353

DB 338 AGAGAGAGATGAAGATACAAATGTGCACACACAGAGAAATGACCCAGTGAGGACACAAAG 397

QY 354 GAGAAATGGCTACTTACAAAGCCTAGGAGAGAGGCTCCGAGAAACACACCCCTACCCAC 413

DB 398 GAGAGGTGGCCACTTACAAAGCCTAGGAGAGAGGCTGGGGGCAACACACCCCTGCCAC 457

QY 414 ACCTTGATGTTGGA-----CTTCATCCTGTAGACGAGGCTCC 452

DB 458 ACCTTGACCTGGACTTCATCCNCCAGATAAAGTCTTCATCCTCCAGATAAAGTCCCTC 517

QY 453 ACCCTCTTCATCAGGTGAAGCCCTTGATTCTGAATATCTCCAAATGCTGGAAGGTAC 512

DB 518 ATCCCTCTTCATCATATGAACCTTCTGACCTGAATCTCTCCAAATGCTGTAATGTAC 577

QY 513 AAAGTGAAGAGACAGACAGACCTCAGGCTGGAAGTTTAAAGAGATAACATCTTCC 572

DB 578 CAAGGTGAAGGACAGACAGACCTCAGGCTGGAAGTTTAAAGAGATAACATCTTCC 637

QY 573 ATTGCTGTGCTCTATCCCTACACACCTATTCCAGTCTTTATTTGGCTTTTGGTTT 632

DB 638 ATTGCCCTGTCCATCCCAACACACACCTGTGCCGCCCTTTATTTGGTCTTTTGTGA 697

QY 633 CCGTGTCCGGGTATAGTGTAGTTGTAATCTGTGTTTACATACAGGATAACATAAACA 692

DB 698 TTCCCTCTGTCCCTGGGATATAAGGGGAACCTTATTAACTCCGGCTTTTATGGTATA 757

QY 693 AAGGTAACAATAAATAAACA 716

DB 758 GGGGATAACCCCTACAAACCAAGA 781

RESULT 4
AQ093103/c
LOCUS
DEFINITION HS_3020_B1.A04.T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3020 Col=7 Row=B, DNA sequence.

ACCESSION AQ093103
VERSION AQ093103.1 GI:3464550
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 391)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL 99380589

MEDLINE Contact: Mahairas GG, Wallace JC, Hood L

COMMENT High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3020 row: B column: 7

Class: BAC ends

High quality sequence stop: 391.

FEATURES Location/Qualifiers

source 1. .391

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=3020 Col=7 Row=B"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 97 a 65 c 45 g 118 t 66 others

ORIGIN

Query Match 18.9%; Score 188.8; DB 17; Length 391;

Best Local Similarity 65.5%; Pred. No. 2.1e-29;

Matches 256; Conservative 0; Mismatches 122; Indels 13; Gaps 2;

QY 506 AAGGTACAAAAGTGAAGACAGACAGACCTCAGGTGAAAAGTTAAAGAGAATAACA 565

DB 378 AAAAAAANNN 319

QY 566 TCTTTCCATTGCTGTCTCTATCCCTACACACACTATCCAGCTTTTATTTGGTCTTTT 625

DB 318 NNANNN 259

QY 626 GTGTTTTCGTGCTCGGTATAGTGTAGTTCTTAATCTGTGTTTACATACAGGATAACA 585

DB 258 GTATTCCCTTTCCTGTGGATAGTGAATTTTAAATTTTGTGTTTACATAGAGATAACA 199

QY 686 TAAACAAAGGTAAACAATAAATAAACAACACAGCAAACTCAACTAATAGTGTGTTG 745

DB 198 TAAAA-----GTAAACAATAAATAAATAAAGAGAAACTCAATAATAGTGTGTTG 144

QY 746 GCATGTCACAGTGAAGACAGGAGAGTCACATAAAATGGAGTGGAACTTTTGAGCTAA 805

DB 143 GGAGGTGACAGTGAAGCGGGAGGTACATAACAATGGAGTGGAGTCTTTTGACTCC 84

QY 806 ATCAATGTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 865

DB 83 ATCAATGTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 32

QY 866 TTAGGTGTTTATCTTCCTGCTGCTCAGTAA 896

DB 31 TTACGTGTTTGTCTCTTGCCTGCTCCTTAA 1

RESULT 5
AA417076/c
LOCUS
DEFINITION zui3b12.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:731711
3', mRNA sequence.

ACCESSION AA417076
VERSION AA417076
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL
COMMENT

Unpublished (1997)
Other_GSSs: CIT978SK-193A5.TP
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. 432

/organism="Homo sapiens"
/db_xref="GDB:5295191"
/db_xref="taxon:9606"
/clone="193A5"
/clone_lib="CIT978SK"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
Caltech Human BAC Library A"

BASE COUNT
ORIGIN

107 a 106 c 89 g 130 t

Query Match

8.9%; Score 89.2; DB 17; Length 432;

Best Local Similarity 69.2%; Pred. No. 1.2e-08;

Matches 137; Conservative 0; Mismatches 58; Indels 3; Gaps 1;

Qy

233 AAAAGTTTAAATGAGGACCTTAGGTGGTCTTAATCCCAATCTAAGTGATGCTCCAT 292

Db

196 AAATATGTGAAGTCAGCTGCTATGGTGGTACTATCTTATATACTATGTCCTTAT 137

Qy

293 GAAAGAGAAATAGGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACACAA 352

Db

136 AAAAGAGAAATAAGA---ACACTGACATGTACAGAGGCAAGACCATGTGAAGACACAG 80

Qy

353 TGAGATGTGCTACTTACAGCTTAGGAGAGAGGCTCCGAGAAACACACCCCTACCCA 412

Db

79 GGAGAAGATGGCCATCTCTAAGCCAAAGGAGAGAGGCTCCGAGAAACACAGACCCCTGTA 20

Qy

413 CACCTTGATGTGGACTT 430

Db

19 CACCTTGATGTAACTT 2

RESULT 8

AG153358

LOCUS

AG153358 Pan troglodytes DNA, clone: RP43-017A19.TJ, genomic survey
sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbegsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.

PRIMERS

Sequencing: TJ

LIBRARY

Vector : pBACE3.6

R.Site 1 : ECORI

R.Site 2 : EORI.

Location/Qualifiers

1. 726

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="RP43-017A19.TJ"

/sex="male"

/cell_type="lymphocytes"

/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 175 a 190 c 178 g 181 t

ORIGIN 2 others

Query Match 8.6%; Score 85.6; DB 17; Length 726;

Best Local Similarity 68.8%; Pred. No. 5.8e-08;

Matches 132; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

Qy 240 TTAATGAGGACCTTAGGTGGTCTTAATCCCAATCTAAGTGATGCTCCATGAAAGAG 299

Db 180 TTGAATGAGGCAATGAGGTGGGCCCTTAATCCCAAGTACTATCATATGAGAAA 239

Qy 300 GAAATAGGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACACAATGAGAT 359

Db 240 GGAGTTAGGACAGAAAG-ACACAGAGAGAGAGGCGCTGTGGGACACAGTGGGAAG 298

Qy 360 GTGGCTACTTACAGCCTTAGGAGAGAGGCTCCGAGAAACACACCCCTACCCACACCTTG 419

Db 299 TTGGCCCTCTACAGCAAGGAGAGAGGCTGCAGAGAAACCAAGCTACCAAGCCCTCG 358

Qy 420 ATGTTGGACTTC 431

Db 359 ACCTTGGACATC 370

RESULT 9

T39269/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

T39269 YAO3B03.r2 Stratagene lung (#937210) Homo sapiens cDNA clone

IMAGE:60365 5', mRNA sequence.

T39269.1 GI:647023

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)

1 (bases 1 to 296)


```

Qy 16 GAAGAAATAGTGCAGAGATCTATTGTATC---TTGGTGACTACACTTAATGTATTGTGTT 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 378 GAAGAAATAGTTCATAGATGGTGTGTCAGTAGTAGTACCATATGTTAATGTATTGTACA 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 73 CTTGA---CTAATACATAGATTTTCAGTGTCTTCACAAACAAAACATGATGGTATGT 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 318 CTTGAAATCGCTGAAGTAGATTTAAGTGTCTTCACCAATAAAGAAATGATAAATATGT 259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 129 GAGGTAATGCATATCAAACTAGCTTGGGTTAACCATTCCACA-ATATGTGTGTATTTC 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 GAGACAAATGCCCAATGTTAACTAGCTTGGTTAGCCATTCACATCATATATATTTAAA 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 188 AAACAGTACCATAATGACAGCAATTTTGTGTACGTTCAATCAAAAAAGTTTAAAAATG 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 ACATGTGTACACCAACAAATACATAATTTTGTGCACTTATATTAAGTAATTAATTAATG 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 248 AGGACCTTAGGTGGTCT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 AAAAATGTTGGTCTATCAT 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
T05554
LOCUS EST03443 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA
DEFINITION clone HFBDC25, mRNA sequence.
ACCESSION T05554
VERSION T05554.1 GI:316704
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE 3,400 expressed sequence tags identify diversity of transcripts
from human brain
JOURNAL Nat. Genet. 4, 256-267 (1993)
MEDLINE 93364420
COMMENT The Institute for Genomic Research
932 Clopper Road, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: mdadams@tigr.org
Seq primer: M13-21.
FEATURES
    source
        1..391
        /organism="Homo sapiens"
        /db_xref="ATCC (inhost):82254"
        /db_xref="taxon:9606"
        /clone="HFBDC25"
        /clone_lib="Fetal brain, Stratagene (cat#936206)"
        /note="Vector: LambdaZAP-II; 17-18 wk gestation, female;
        oligo-dT + random primed cDNA synthesis; lambdaZAP-II
        vector, 1.0kb average inser size."
BASE COUNT 119 a 85 c 92 g 91 t 4 others
ORIGIN
Query Match 8.2%; Score 82.4; DB 14; Length 391;
Best Local Similarity 68.8%; Pred. No. 3.1e-07;
Matches 132; Conservative 0; Mismatches 51; Indels 9; Gaps 1;

Qy 240 TTAATAATGAGGACCTTAGGTGGTCCCTAATCCATCTAAGTGATCTCTCCATGAAGAG 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74 TTAATAATGAAGACATTAGGTGGGCGCCTTAATCCATCTGACTGGTCTCTCTG----- 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 GAAATGAAGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACACATGAGAA 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 --AAGAGAGGAGATGAGGACACATGCGAGAGGCATGACCACATGAGGACACAGGAGAG 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 360 GTGGCTACTTACAGCCTTAGGAGAGAGGCGCTCCGAGAAACACACACCCCTACCCACACCTTG 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 185 GTGGCCATCCCAATCAAGGAGTGAGCCCTCGGAGGAGACAGCACTTACCACACCCCTG 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 420 ATGTTGGACTTC 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 ATCCTGGACTTC 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
A0468873
LOCUS HS_5140_B2_E01_T7A RPCI-11 Human Male BAC library GSS 23-APR-1999
DEFINITION genomic clone Plate-716 Col=2 Row=J, DNA sequence.
ACCESSION A0468873
VERSION A0468873.1 GI:4645969
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 550)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 716 row: J column: 2
Seq primer: T7
Class: BAC ends
High quality sequence stop: 550.
FEATURES
    source
        1..550
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="Plate-716 Col=2 Row=J"
        /clone_lib="RPCI-11 Human Male BAC Library"
        /sex="male"
        /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
        Male blood DNA was isolated from one randomly chosen donor
        and partially digested with a combination of EcoRI and
        EcoRI Methylase. Size selected DNA was cloned into the
        pBACe3.6 vector at EcoRI sites"
BASE COUNT 158 a 111 c 129 g 144 t 8 others
ORIGIN
Query Match 8.2%; Score 82.4; DB 17; Length 550;
Best Local Similarity 67.4%; Pred. No. 2.9e-07;
Matches 130; Conservative 0; Mismatches 62; Indels 1; Gaps 1;

Qy 240 TTAATAATGAGGACCTTAGGTGGTCCCTAATCCATCTAAGTGATCTCTCCATGA-AAGA 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 TTAAGTGAAGTTATTGTAGTGGGCGCCTTAATCCAAATGACATGATCATTTACAAGATGA 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 299 GAAATATAAGGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACCAATGAGAA 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 GGAATTTAGACAGACACACACAGTTACAGAGGGAAGACCGTGAAGACACAAGGGGAT 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 359 TGTGCTACTTACAAGCCTTAGGAGAGAGGCGCTCCGAGAAACACACACCCCTACCCACACCTT 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 308 GGTGGCCATGTACAAGCCTTANGACAGAGGCTTCGGGAGGAGCAACCCAGCCACACCTT 367
QY 419 GATGTTGGACTTC 431
Db 368 GATTCAGACTTC 380

Search completed: December 24, 2002, 21:53:54
Job time : 1117.8 secs

THIS PAGE BLANK (USPTO)